

A and B. If the degree of overlap is large (figure 10.1a) it will be difficult to distinguish A from B by sampling, but it will also not be very important. However, if the degree of overlap is small (figure 10.1b) it will be valuable to distinguish A from B, but it will also be easier. The suggestion is that sampling will be most important at some intermediate level of overlap. I call this the self moderation hypothesis (because the value of sampling is presumed to have a natural limit), and I use the values of optimal behaviour and sampling defined in the previous section to consider it.

Contrary to the self-moderation hypothesis several workers in patch sampling theory have found that departures (in optimal behaviour and in the pay-offs due to optimal behaviour) from non-sampling models are greatest when overlap is smallest, (Green 1980, Iwasa et al 1981, McNamara 1982; notice that in this case decreasing overlap is the same as what has been more generally called increasing contagion or variance in the mixing distribution of patch sub-types). Both results can be shown to be correct in terms of the definitions of the previous section. The difference is the set Ω against which optimal behaviour is compared. I illustrate this with a simple sampling problem, which reappears in two subsequent sections.

The model: Imagine that a forager encounters a class of patches which consists of two indistinguishable sub-types, 1 and 2. All patches contain two prey items, but the forager cannot assess their value until it eats them. In both patch sub-types the values of prey items are random variables distributed according to a normal distribution. For convenience, I refer to prey value as "size," although in this example it will often be negative. It takes one unit

of time to eat each prey item. Patches of sub-type 1 provide two prey items whose sizes are determined independently according to a normal distribution with mean ℓ and variance σ^2 . Patches of sub-type 2 provide items drawn from a normal with mean $-\ell$ and variance σ^2 . For convenience, I will refer to this type of patch simply as an " ℓ -system of patches." Notice that this system is closely analogous to that presented in figure 10.1 with ℓ controlling the degree of overlap. If ℓ is zero, overlap is total; and if ℓ is large overlap is small.

In this case sampling behaviour is determined by the choice of a critical prey size c , such that if the first prey eaten is smaller than c the forager abandons the patch and moves on to the next, after some travel time $1/\lambda_S$. The only possible form for a non-sampling strategy is to choose how many prey should be taken from each patch. In this case, where there are no alternatives to the ℓ -system, a non-sampling forager should always take both prey in the patch. In terms of the previous section, the universal set of strategies is simply the set of all possible critical prey sizes for $c = -\infty$ to $c = \infty$; $c = -\infty$ corresponds to the non-sampling strategy "always take both." Given a c -rule the rate of energy gain per time unit can be found. Let c^* be the c which provides the maximum rate R_S^* .

$$R_S^* = \frac{\lambda_S \left(\frac{\ell}{2}\right) \left\{ \Phi\left(\frac{c^* + \ell}{\sigma}\right) - \Phi\left(\frac{c^* - \ell}{\sigma}\right) \right\}}{1 + \lambda_S \left(\frac{1}{2}\right) \left\{ 4 - \Phi\left(\frac{c^* + \ell}{\sigma}\right) - \Phi\left(\frac{c^* - \ell}{\sigma}\right) \right\}}, \quad (10.3)$$

where Φ is the cumulative distribution function of a standard normal distribution. The rate of gain from this system for any non-sampling strategy must be zero. So expression (10.3) is also the value of sampling, $W_S = R_S^*$. When there is complete overlap $\ell = 0$ and

$R_S^* = 0$ for any sampling rule. The value of sampling, as we would expect, is therefore zero when there is total overlap. However, as ℓ becomes large (larger than about 2σ) it is possible to choose c^* so that $\Phi\{(c^*-\ell)/\sigma\} \approx 0$ and $\Phi\{(c^*+\ell)/\sigma\} \approx 1$. That is, for large ℓ

$$W_S = R_S^* \approx \frac{\lambda_S \ell}{2 + 3\lambda_S} \quad (10.4)$$

Therefore, the value of sampling obviously increases monotonically and without limit as ℓ increases, in accordance with the theoretical work of Green (1980), Iwasa et al (1981) and McNamara (1982); this is in contradiction to the seemingly sensible self-moderation hypothesis.

However, consider the value of optimal behaviour with respect to a simple sampling rule of thumb. Imagine, for example, that the forager's faculties are limited so that it can only distinguish positive from negative prey values; i.e. whether it makes a net gain on a prey item, or a net loss. The forager should adopt a rule which tells it to leave patches in which it samples negative prey values and to stay otherwise; that is, it sets $c=0$. The value of optimal behaviour with respect to this zero rule $W(c=0)$ is:

$$W(c=0) = R_S^* - \frac{\lambda_S \ell \{1 - 2\Phi(-\frac{\ell}{\sigma})\}}{2 + 3\lambda_S} \quad (10.5)$$

When there is complete overlap $W(c=0) = 0$, and when ℓ is large $\Phi(-\ell/\sigma) \approx 0$ so that as ℓ approaches infinity, $W(c=0)$ approaches 0. In figure 10.2 I graph the value of optimal behaviour with respect to this sampling rule of thumb, and find that it behaves exactly as we would expect according to the self-moderation hypothesis.

I conclude that Green (1980), Iwasa et al (1981) and McNamara (1982) are quite right to assert that the value of sampling is greatest

Figure 10.2 The Value of Optimal Behaviour with respect to a simple sampling rule of thumb.

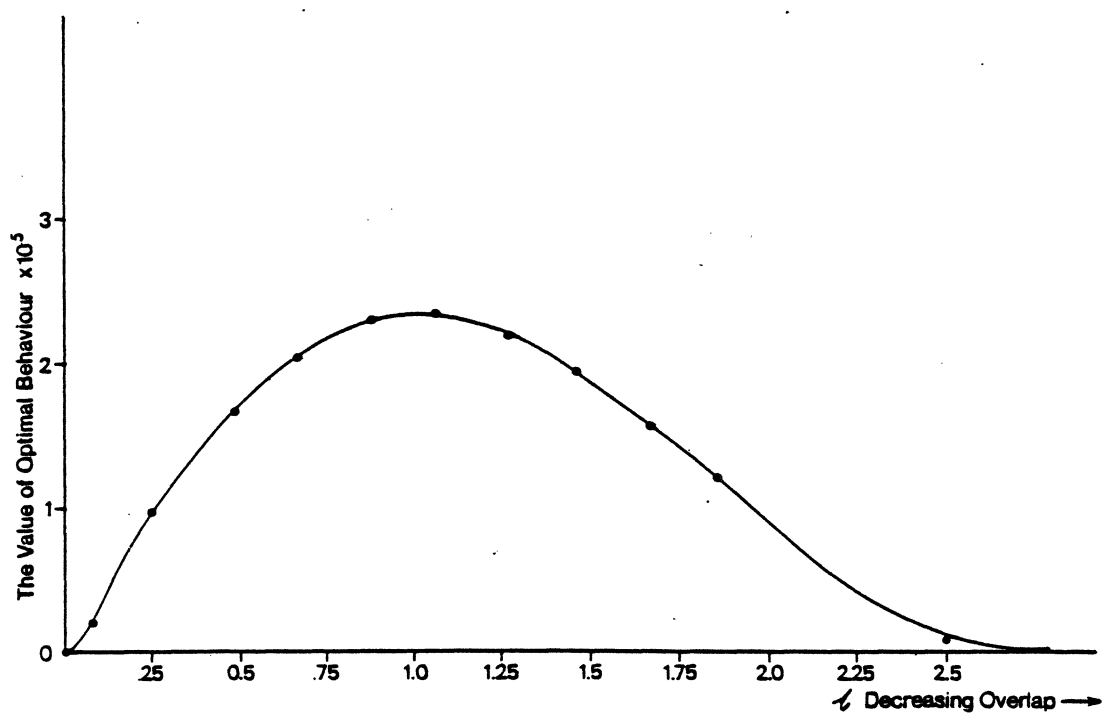


Figure 10.2 The figure illustrates the self-moderation hypothesis, by showing the value of optimal behaviour (compared to a simple sampling rule) as a function of decreasing overlap (increasing l). The value of optimal behaviour is highest at intermediate levels of overlap.

when overlap is smallest; but this is also the condition in which simpler sampling rules of thumb will be most effective, as the self moderation principle would suggest.

10.4 The Importance of Alternatives.

In an interesting way, perfect recognition and perfect ignorance are theoretically similar. Suppose for example that a forager knows that one of the n prey types it encounters is actually composed of two indistinguishable sub-types, a and b. Each sub-type having associated with it mean net energy gains and handling times, E_a , E_b , h_a , h_b respectively. Imagine that the forager can distinguish a's from b's only after they have been eaten. Then as long as a's and b's are encountered independently of one another, the forager should make its decision about whether or not to exploit the a/b-system upon encounter exactly as if it were an unambiguous prey type with profitability $(p_1 E_a + p_2 E_b) / (p_1 h_a + p_2 h_b)$, where $p_1 = P(a \mid \text{an a or b has been encountered})$ and $p_2 = 1 - p_1$. In this case, it is actually best to act as if the mean were known with certainty (cf. Raiffa 1968); i.e. to act as if there were no sub-types. Notice that this depends critically on the two underscored statements. As the analysis of chapters eight and nine shows, if an encounter with a sub-type a item tells the forager that the next encounter is more likely to be with another a, then the forager should consider the value of information gained from sampling the system. Perhaps more interestingly, however, is the importance of alternatives to the a/b-system. If the a/b-system is the only prey type then the mean profitability has no meaning since the forager has no choice but to always take a/b's when encountered. A second goal for this analysis

of the value of sampling is to consider the what factors might make such non-sampling, mean-equivalent rules more attractive to a forager.

In this section I consider only a special type of alternative patch. Imagine that in addition to the ℓ -system of patches introduced in the previous section, the forager also encounters and recognizes as distinct a patch type A (for alternative). Type A patches are not divided into sub-types, (i.e. they are unambiguous). If the forager enters patches of type A it gains a units of energy and spends one unit of time. The universe of possible strategies is composed of all possible choices of the variables, p_a , p_s , c , n_s , (p_a , $p_s \in \{0,1\}$, $c \in \{-\infty, \infty\}$, and $n_s \in \{0,1,2\}$). p_a is the probability that an A patch will be entered and exploited upon encounter. p_s is the probability that the ℓ -system will be exploited upon encounter. c is the critical size of the first prey observed in an ℓ -patch as described in the previous section. n_s is the number of prey taken from an ℓ -patch when encountered.

All possible non-sampling strategies may be represented by a choice of p_a , p_s and n_s . However, in this case, it is easy to show that n_s should either be 0 or 2, that is either the minimum or maximum possible (a result of the zero-one rule of optimal diet theory; see e.g. Charnov 1976b). This rather trivial result will be important in the next section, so I call it the extreme residence time rule. Because of this rule we need only consider the universal set of strategies to be composed of p_a , p_s and c ; and the set of all non-sampling strategies as defined by p_a , p_s (i.e. $n_s = 2$).

Let G_a and H_a be the random variables which describe the gain per encounter with patches of type A and the time spent per encounter

with A's respectively. Similarly define G_S and H_S for encounters with the ℓ -system. Then the rate of energy gain is

$$R = \frac{\lambda_a E(G_a) + \lambda_s E(G_s)}{1 + \lambda_a E(H_a) + \lambda_s E(H_s)} \quad (10.6)$$

where λ_a and λ_s are encounter rates with the alternative and ℓ -systems respectively.

$$\begin{aligned} E(G_s) &= \ell p_s \left(\frac{1}{2}\right) \left\{ \Phi\left(\frac{c + \ell}{\sigma}\right) - \Phi\left(\frac{c - \ell}{\sigma}\right) \right\} \\ E(H_s) &= p_s \left(\frac{1}{2}\right) \left\{ 4 - \Phi\left(\frac{c + \ell}{\sigma}\right) - \Phi\left(\frac{c - \ell}{\sigma}\right) \right\} \end{aligned} \quad (10.7)$$

$$E(G_a) = p_a a \quad E(H_a) = p_a$$

The best strategy: Given that the values of p_a and p_s are specified the optimal value of c (c^*) may be found by solving the implicit non-linear equation:

$$R^* = \frac{\ell \left\{ \Phi\left(\frac{c^* - \ell}{\sigma}\right) - \Phi\left(\frac{c^* + \ell}{\sigma}\right) \right\}}{\Phi\left(\frac{c^* + \ell}{\sigma}\right) + \Phi\left(\frac{c^* - \ell}{\sigma}\right)} \quad (10.8)$$

where $\phi(z)$ is the probability distribution function of a standard (mean zero, variance one) normal distribution. This is an intuitively appealing condition since it says that a forager should leave an ℓ -patch if its expectation of the size of the second prey item, given the observed size of the first, is greater than the rate of gain in the environment.

The best non-sampling strategy: Since the expected gain from any non-sampling rule applied to the ℓ -system is zero, the best non-sampling rule is easily found:

1. If $a \geq 0$ Choose A patches only
(set $p_a = 1$, $p_s = 0$)

$$R_N = \frac{\lambda_a a}{1 + \lambda_a} \quad (10.9)$$

2. If $a < 0$ Choose ℓ -patches only
(set $p_a = 0$, $p_s = 1$)

$$R_N = 0$$

where R_N is the rate due to the best non-sampling strategy.

Taken together these results suggest the following algorithm for computing the value of sampling, (let R^* be the optimal rate of gain):

Algorithm 10.1

1. Set $p_s = 1$
 - 1a. Set $p_a = 0$
Compute c^* , R^*
Set $c_0^* = c^*$, $R_0^* = R^*$
 - 1b. Set $p_a = 1$
Compute c^* , R^*
Set $c_1^* = c^*$, $R_1^* = R^*$
2. Set $p_s = 0$, $p_a = 1$
Compute R_N
Set $R_N = \max\{R_N, 0\}$
3. Set $R^* = \max\{R_0^*, R_1^*, R_N\}$
4. Set $W_S = R^* - R_N$.

The computations in step 1 are solutions to the implicit equation (10.8). I perform this calculation numerically using the method of bisections for finding the roots of non-linear equations, (see Conte and de Boor 1965).

By an argument similar to that used to derive expression 10.7 in the previous section, if ℓ is large the value of sampling is given by

$$W_S \approx \max\left\{\frac{\lambda_a a + \lambda_s (\ell/2)}{1 + \lambda_a + \lambda_s (3/2)}, \frac{\lambda_s (\ell/2)}{1 + \lambda_s (3/2)}\right\} - \max\left\{\frac{\lambda_a a}{1 + \lambda_a}, 0\right\} \quad (10.10)$$

which is, in any case, an increasing linear function of ℓ . In the numerical results in this and the next section, linearity in ℓ is shown to hold at quite small values of ℓ , (figures 10.3 and 10.4).

Notice that if $a < 0$ the alternative patch type should always be ignored, and the value of sampling (W_S) is unaffected by the presence

of the alternative. However, if $a > 0$ (i.e. greater than the mean value of the λ -system) then the value of sampling is always reduced by increasing λ_a . λ_a is a measure of the abundance in the forager's environment of known alternatives. If $\lambda_a = 0$ the forager has no alternative to the λ -system. The effect of increasing λ_a can be shown by calculating the value of sampling at several values of λ_a ($=0, 0.1, 1$) and over a large range of λ 's using algorithm 10.1. The results are shown in figure 10.3.

The results of these calculations suggest that the value of sampling when there are no recognizable and unambiguous alternatives to the sampled system of patches is always greater than or the same as the value of sampling when there are some known alternatives. Many theoretical analyses of patch sampling behaviour have only considered how the forager should treat a single system of indistinguishable patch sub-types, (Oaten 1977, Green 1980, Iwasa et al 1981). This result suggests that this is an extreme situation; when known and recognizable alternative patch types are available and abundant, sampling procedures may be significantly less valuable. It may be unfair to compare sampling to non-sampling "mean" based strategies in cases where there are no recognizable alternatives because such "mean" strategies can only have meaning when there are alternatives to compare.

10.5 Patch depression and patch sampling.

Green (1980) has severely criticized Charnov's (1976) perfect patch recognition model of optimal patch residence time, by presenting his own modification of Oaten's (1977) patch sampling model. Charnov's model assumes that the rate of energy gain within a patch decreases as the patch is exploited; in mathematical terms, the gain function is

Figure 10.3 The effect of Recognized Alternatives on the value of Sampling.

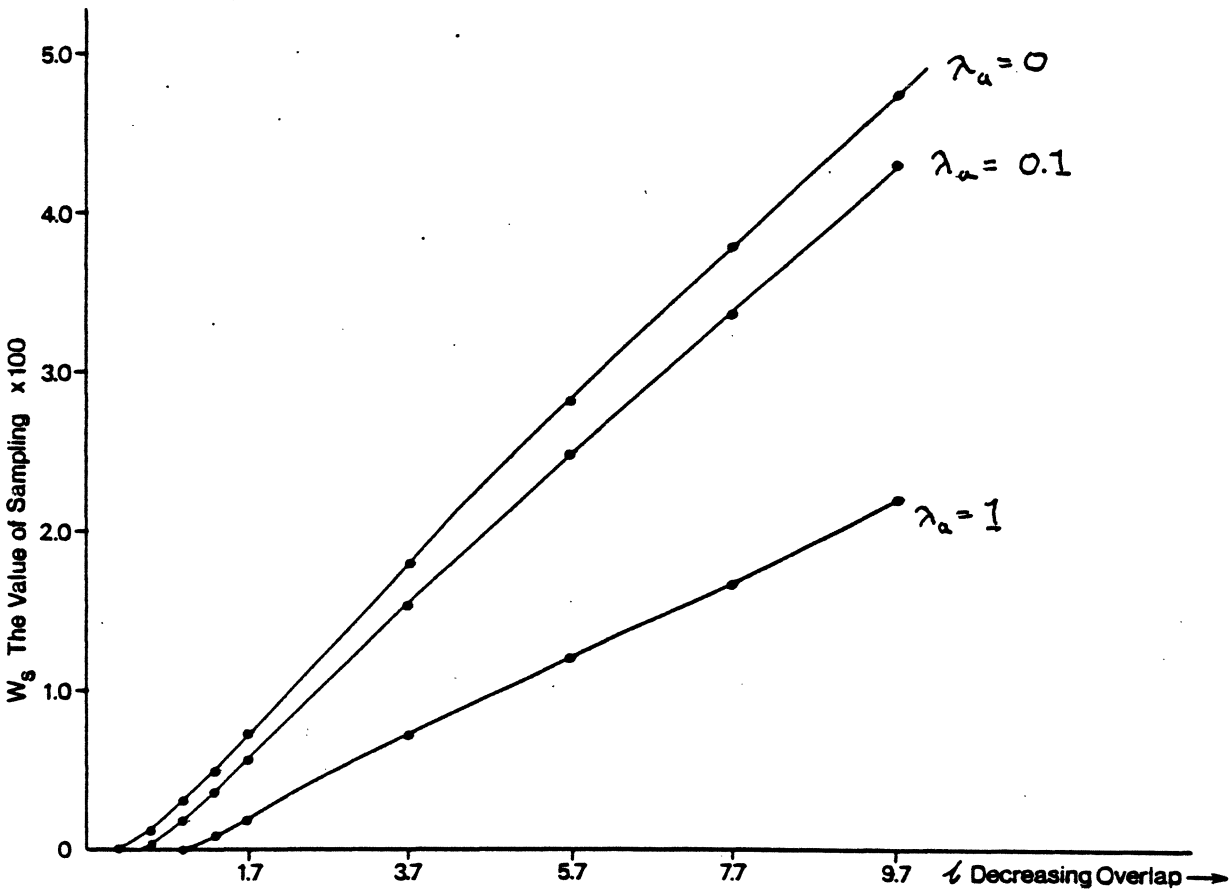


Figure 10.3 The figure shows three plots of the value of sampling as a function of decreasing overlap. The plots have different λ_a 's the encounter rate with alternatives. $\lambda_a=0$, $\lambda_a=0.1$, $\lambda_a=1$. As λ_a increases the value of sampling may decrease as shown.

negatively accelerated. This negative acceleration may be due to several factors, such as random search of a depleting patch, or decreasing ability to capture prey (as in Orians and Pearson's 1979 "central place" version of Charnov's result). Of all Charnov's assumptions this is perhaps the least questionable (see Charnov et al 1976 for a discussion); most biologists agree that some negative acceleration or depression (in the terminology of Charnov et al 1976) is natural.

However, Green's (1980) model assumes a special sort of patch in which search is systematic and in which there is no patch depression. This means that the validity of Green's model as a criticism of Charnov's hinges upon a sort of independence of patch depression and the value of sampling. It is unlikely that this is the case. To see this consider Green's (MS) statement that Charnov's model is "like telling a gambler that he should stop playing just before he starts losing." This analogy is clearly wrong for a forager which experiences patch depression; because it is a gambler who knows that, at best, things are getting worse. This is a significant and useful piece of information. More importantly, if depression is at all general it is a piece of information animals will often have.

In this section, I consider how patch depression and the value of sampling interact by building depression into the λ -system of patches discussed in the previous two sections. In these simple two prey patches, depression implies that the mean size of the second prey item is less than the mean of the first. Let $X_{i,j}$ be the random variable representing the size of the j th item "observed" from a patch of type i . Let the variable d represent depression.

$$\begin{aligned} E(X_{1,1}) &= l + d & E(X_{2,1}) &= -l + d \\ E(X_{1,2}) &= l - d & E(X_{2,2}) &= -l - d \end{aligned}$$

As above, the X 's are normally distributed with equal variances, σ^2 . Assume that there are no alternatives to the l -system; so that the best sampling strategy is determined by the choice of the critical value c .

The best sampling strategy. The equations describing this case are only slightly different from those of section 10.4:

$$\begin{aligned} E(G_S) &= \left(\frac{1}{2}\right) \left\{ \phi\left(\frac{c-l-d}{\sigma}\right)(d-l) + \phi\left(\frac{c+l-d}{\sigma}\right)(d+l) \right\} \\ E(H_S) &= \left(\frac{1}{2}\right) \left\{ 4 - \phi\left(\frac{c+l-d}{\sigma}\right) - \phi\left(\frac{c-l-d}{\sigma}\right) \right\} \end{aligned} \quad (10.11)$$

Since there are no alternatives:

$$R = \frac{\lambda_S E(G_S)}{1 + \lambda_S E(H_S)} \quad (10.12)$$

The condition for optimization is to choose c (c^*) to satisfy:

$$R = \frac{(\lambda-d)\phi\left(\frac{c^*-l-d}{\sigma}\right) - (\lambda+d)\phi\left(\frac{c^*-d+l}{\sigma}\right)}{\phi\left(\frac{c^*-l-d}{\sigma}\right) + \phi\left(\frac{c^*-d+l}{\sigma}\right)} \quad (10.13)$$

The best non-sampling strategy. In section 10.4 I described the extreme residence time rule, which stated that the optimal non-sampling strategy was either (i) to not enter the l -system of patches or (ii) to always enter and eat the maximum number of prey items, (i.e. 2). Perhaps the most important distinction between the depression and no depression cases is that the extreme variance time rule does not hold when there is patch depression. Since there are no alternative patch

types, a non-sampling forager only has the choice of observing one prey size per patch, or observing both (i.e setting $n_s = 1$ or 2). In this case setting $n_s = 1$ is always the optimal behaviour for a non-sampling forager; since $R_{n_s=2} = 0$ and $R_{n_s=1} = \lambda_s d / (1 + \lambda_s)$.

Notice that in the patch depression case there are conditions of ℓ and d in which the non-sampling strategy (set $n_\ell = 1$) is as good as the best sampling strategy (c^* ; where $c^* = \infty$). In these conditions (large d , small ℓ) the value of sampling is obviously zero. This happens because, when patches depress, the forager has valuable information about current patch quality based wholly on the time since entering the patch, irrespective of sampling.

This suggests a modification of algorithm 10.1 to find the value of sampling in this case.

Algorithm 10.2

1. Calculate c^* from equation (10.16).
2. Calculate $R(c^*)$.
3. Calculate $R_N = \lambda_s d / (1 + \lambda_s)$.
4. $R_{\max} = \max\{R(c^*), R_N\}$.
5. $W_S = R_{\max} - R_N$.

I have calculated algorithm 10.2 for four levels of depression, and at each level of depression I have found W_S over a large range of ℓ 's. In figure 10.4 I present the results of these calculations. The figure shows the value of sampling W_S as a function of ℓ for four levels of depression. As in section 10.3 at large ℓ , c can always be chosen so the $\Phi[(c-\mu-d)/\sigma] \approx 0$ and $\Phi[(c-d+\mu)/\sigma] \approx 1$. So that at large ℓ :

$$W_S \approx \frac{\lambda_s \{(d+\ell)/2\}}{1 + \lambda_s (3/2)} - \frac{\lambda_s d}{1 + \lambda_s} \quad (10.14)$$

which is a linear equation in ℓ , with slope $\alpha_1 = \lambda_s / (2 + 3\lambda_s)$ and

Figure 10.4 The Value of sampling at different levels of patch depression, d.

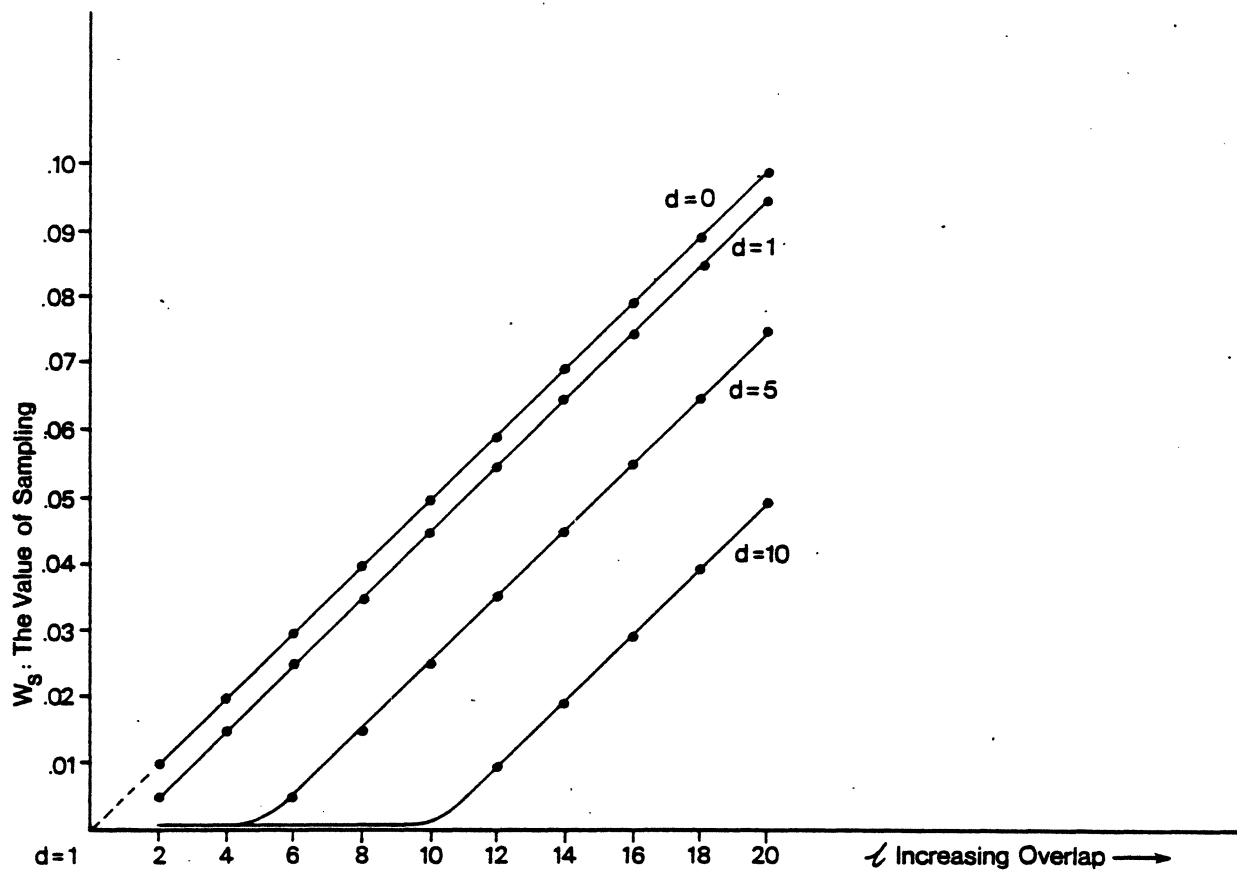


Figure 10.4 This figure shows four plots of the value of sampling as a function of decreasing overlap, each a different level of patch depression d, d=0,d=1,d=5,d=10. The figure shows that increasing depression decreases the value of sampling.

intercept

$$\alpha_2 = \frac{-\lambda_s^2 d}{2(2 + 3\lambda_s)(1 + \lambda_s)}$$

Figure 10.4 shows that linearity obtains very quickly. This allows us to write an approximate formula for w_S .

$$w_S \approx \begin{cases} 0, & \ell < d\left(\frac{1 + 2\lambda_s}{1 + \lambda_s}\right) \\ \alpha_1 \ell + \alpha_2, & \text{otherwise.} \end{cases} \quad (10.15)$$

The clear result of this analysis is that at least in this case, patch depression reduces the value of sampling. Of course, the presence of depression does not mean that patch sampling is unimportant; it may be quite important when the distribution of patch types is very contagious, (large ℓ). When ℓ is large, however, cheaper sampling strategies may also be very good. This result does not invalidate Green's fascinating model, especially since there are some important natural patch types without depression, (see the next section), but it does mean that Green's model cannot be fairly applied as a criticism of Charnov's (1976) marginal value theorem, in which depression is the key element.

10.6 Partial patch recognition.

In section 10.4 I argued that previous theoretical work on the problem of patch sampling has considered cases in which the value of sampling is unnaturally inflated; because they do not consider the possible presence of recognizable and unambiguous alternative patch types, or patch depression. This does not mean that patch sampling will be unimportant when there are recognized alternative patch types. It does suggest that it will be important to add the recognition of

alternative patch types to patch sampling models. In this section I consider a simple and natural model of this kind. The model shows that by considering patch sampling and patch recognition together we can explain some previously puzzling foraging phenomena.

The Model. Consider a simple kind of forager. It encounters discrete but homogenous patches of differing sizes. Let z_i be the size of the i^{th} patch type. The forager hunts (within patches) for immobile prey items systematically; that is, it hunts without retracing its steps within a patch. Irrespective of patch size, patches are of two types: empty and full. Empty patches contain no prey. Full patches are filled with prey at random according to a Poisson process with parameter θ ; so that the mean number of prey per patch equals $\theta \cdot (\text{patch size})$. The probability that a given patch encountered is full is p , [therefore $1-p=P(\text{empty})$]. τ is the mean travel time between patches.

I assume that the forager can recognize patch size instantly upon encounter, but that it cannot recognize patch type, (i.e. empty or full). The simplicity of this model is due to the form of the rule for optimal behaviour. If the forager encounters a prey item, it knows the patch is full, and it should finish searching it exhaustively. Therefore, the forager should use a simple giving-up time rule for each patch size, z_i . Such a rule has the form:

1. If no prey item has been encountered before t_i leave the patch and go to another.
2. If at least one prey item has been encountered before t_i , stay and finish searching the patch exhaustively.

Thus a complete specification of the forager's strategy consists of a list of one giving-up time (t_i) for each patch size z_i .

Assume that patch size z_i occurs in proportion δ_i , $\{\sum \delta_i = 1\}$. For convenience, I measure size in units of time taken to search a patch systematically and exhaustively. Finally, I assume that by some other optimization procedure the forager has already decided to exploit all patches greater than a certain size, and I consider only these to have non-zero δ_i 's.

The first step in the model is to express the rate of energy gain (which I wish to maximize) as a function of the list of giving-up times $\{t_1, t_2, \dots, t_k\}$. Let N_{z_i} be the random variable which represents the number of prey obtained from a patch of size z_i . Clearly,

$$E(N_{z_i}) = E(N_{z_i} | \text{full})p. \quad (10.16)$$

Let T_1 be the random variable that represents the time from entering the patch until the first prey item is encountered. In full patches T_1 follows a negative exponential distribution with rate parameter, θ . From the the properties of conditional expectations I find:

$$E(N_{z_i} | \text{full}) = E(N_{z_i} | T_1 \leq t_i)P(T_1 \leq t_i | \text{full}). \quad (10.17)$$

From the properties of the negative exponential distribution.

$$\theta z_i = E(N_{z_i} | T_1 \leq t_i)P(T_1 \leq t_i | \text{full}) + \theta(z_i - t_i)P(T_1 > t_i | \text{full}) \quad (10.18a)$$

or by rearrangement and substitution

$$E(N_{z_i} | T_1 \leq t_i)P(T_1 \leq t_i | \text{full}) = \theta z_i - \theta(z_i - t_i)\exp\{-\theta t_i\} \quad (10.17b)$$

Therefore

$$E(N_{z_i}) = p\theta[z_i - (z_i - t_i)\exp\{-\theta t_i\}] \quad (10.18)$$

By a similar argument, let T_{z_i} be the random variable describing

the time spent in patch of size z_i .

$$\begin{aligned} E(T_{z_i} | \text{full}) &= z_i - (z_i - t_i) \exp\{-\theta t_i\}, \\ E(T_{z_i} | \text{empty}) &= t_i. \end{aligned} \quad (10.19)$$

After some algebra I find that

$$E(T_{z_i}) = p(z_i - t_i)(1 - \exp\{-\theta t_i\}) \quad (10.20)$$

The rate of prey gain per unit time can now be expressed as a function of the t_i 's, as

$$R = \frac{\theta \sum_{i=1}^k \delta_i \{z_i - (z_i - t_i) \exp\{-\theta t_i\}\}}{\tau + \sum_{i=1}^k \delta_i \{p(z_i - t_i)(1 - \exp\{-\theta t_i\})\}} \quad (10.22)$$

By differentiation I find that R is maximized by choosing the t_i 's so that

$$R = \frac{p \exp\{-\theta t_i\} \{(z_i - t_i)\theta + 1\}}{p\{\theta \exp\{-\theta t_i\}(z_i - t_i) - (1 - \exp\{-\theta t_i\})\} + 1} \quad (10.23)$$

or equivalently,

$$\{\theta(z_i - t_i) + 1\} \exp\{-\theta t_i\} = \frac{R(1 - p)}{p(\theta - R)} \quad (10.24)$$

The right hand term is the same for all patch sizes, so the (10.24) specifies the form of a relationship which must hold between patch size and giving-up time. The shape of this curve is plotted in figure 10.5. This shows that in bigger patches the forager should be more tenacious. Notice that at any time at which nothing has yet been found the forager's subjective assessment of the probability that the patch is full is the same regardless of patch size. The forager stays longer in big patches because the potential value of finding out that a big

Figure 10.5 Predicted Optimal Giving-up time (GUT) as a function of patch size.

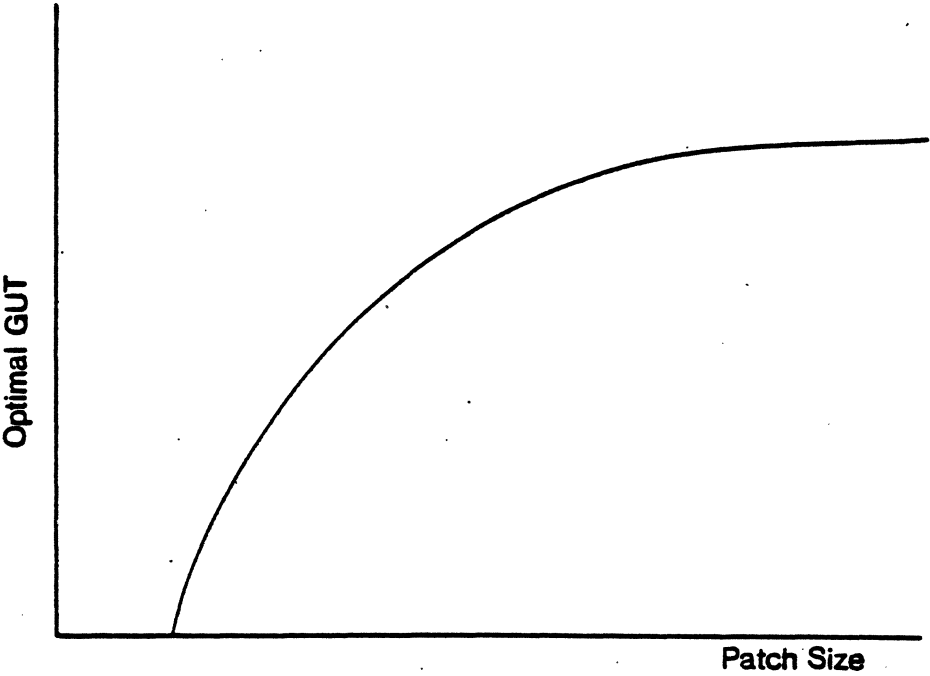


Figure 10.5 This figure shows the prediction of the partial patch recognition model of section 10.6 that patch tenacity should increase with patch size.

patch is full is greater than the value of finding that a small patch is full, (see McNamara 1982 for an interesting development of the idea of the potential value of information). This is a case in which a forager should use information it already has (about patch size) as well as information gained while in the patch to make an optimal decision. It is a case which cannot be modelled by the Charnov's (1976) perfect recognition model, nor by the Oaten (1977) and Green (1980) family of models in which all patches are recognized as identical before sampling begins. This model is similar to Green's (1980), in that there is no patch depression, since within patch search is systematic.

Of course, this model makes many simplifying assumptions, and it will have limited application to the field. However, nectar-feeding animals feeding on linear inflorescences face a surprisingly similar problem. Pyke (1982) has studied bumblebees (Bombus flavifrons) foraging of vertical inflorescences of monkshood, (Aconitum columbianum). The bumblebees nearly always start at the bottom flower of an inflorescence and move systematically upwards. By mid-morning most, but not all, flowers have been visited and many are empty or nearly so. This creates what is known in the bee foraging literature as a bonanza-blank nectar distribution (cf. Zimmerman 1979). Inflorescences vary greatly in size, having between one and fourteen flowers. Given the well documented abilities of bees to recognize flower shape and colour (e.g. Heinrich 1979), there is no reason to suppose that the bees should have any difficulty assessing the size of inflorescences.

Pyke (1982) has measured patch tenacity by considering the probability that a bumblebee will leave an inflorescence after visiting

a single flower. He has also demonstrated that the distribution of nectar rewards per flower is not different for large (many flower) and small (few flower) inflorescences. The model developed here predicts that the bees should be more tenacious on larger inflorescences. This is exactly what Pyke finds, figure 10.6. Pyke correctly points out that neither Charnov's nor Oaten's apparently alternative models account for this effect. This explanation requires both patch recognition (not in Oaten's model) and sampling, (not in Charnov's model).

There may be other explanations for this effect but this one should be testable. A test would consist of experimentally altering the potential value of information, for example, by altering the probability that a patch is full, p . This should affect the degree of this tenacity effect, from non-existent when there is no potential information (e.g. $p=1$ or $p=0$) to strong when there is potential information, ($p=.5$).

10.7 Discussion and summary.

In this chapter I propose a definition for the value of sampling, in the context of models of patch sampling. The definition proposed may also be applied to the value of optimal behaviour with respect to the best behaviour chosen from a subset of strategies, or rules of thumb. I apply this definition to three problems. Firstly, I consider the idea that sampling is most valuable at limited levels of contagion in the mixing distribution of patch types. I find that as I have defined the value of sampling this is not the case; as the difference between two indistinguishable patch sub-types increases, the value of sampling always increases. However, the self-moderation hypothesis is

Figure 10.6 Pyke's (1982) evidence of patch tenacity. Bumblebees feeding on Monkshood.

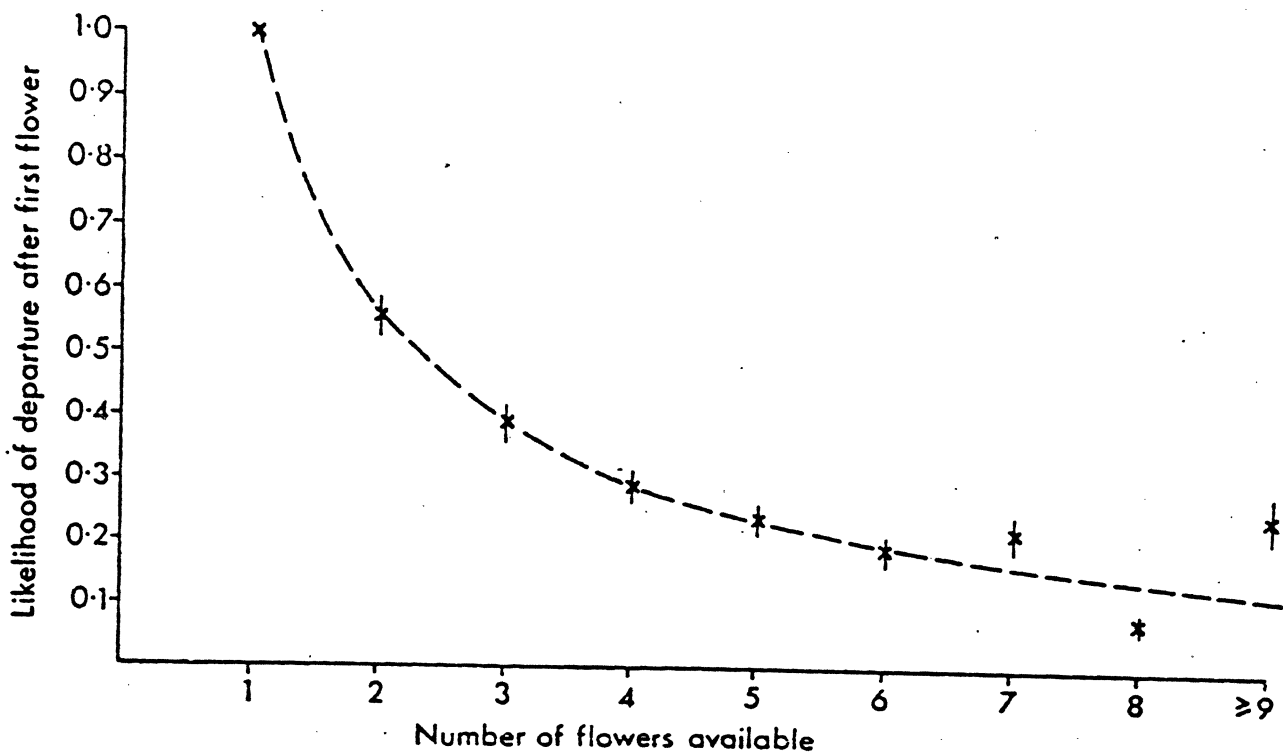


Figure 10.6 This figure shows Pyke's data on patch tenacity. The number of flowers available (patch size) is on the abscissa, the relative frequencies of departures after the first flower is on the ordinate. The monotonic decrease shown is consistent with partial patch recognition (as is shown in section 10.6), but not consistent with perfect recognition (Charnov 1976) or perfect ignorance (Oaten 1977).

applicable for simple, but not optimal, sampling rules of thumb; i.e. the value of optimal behaviour with respect to such simple rules of thumb is highest at intermediate levels of contagion. Two papers have shown that patch sampling makes the most difference when the degree of overlap is lowest, (Green 1980, Iwasa et al 1981). Since this is also the case where sampling rules of thumb are most effective, this may mean that rules of thumb will be very important in sampling situations. Secondly, many models of patch sampling (Oaten 1977, Green 1980, Iwasa et al 1981) have assumed that the forager faces a monolithic system of indistinguishable patch sub-types; that is, that the forager recognizes nothing about patches before entering them. I have modelled a simple system of indistinguishable patches and considered the effect on the value of sampling of adding a simple type of alternative and recognizable patch type; i.e. one that is of known and unambiguous value. I find that the value of sampling in the presence of such alternatives is either the same or less than the value of sampling without them. If foraging animals face such unambiguous and recognizable patches in nature then conventional models of patch sampling will tend to over-value sampling behaviour. It is important to realize that this argument comes from considering a rather special type of alternative patch type, and it does not mean that sampling will necessarily be unimportant when many patch types are distinguishable, (cf. section 10.6). Instead, this result means that the addition of recognizable alternatives can lower the value of sampling; and that a simple class of alternatives always will.

Thirdly, Green (1980) has presented a model of patch sampling as a basis of criticism of Charnov's (1976) perfect knowledge model of

optimal use of depressing patches. However, Green's model assumes ^{there} that ~~is~~ no patch depression. I use the value of sampling defined in this chapter to consider whether the presence and degree of patch depression affects the value of patch sampling. I find, in a simple model, that the value of sampling is always reduced by adding patch depression to the model. I conclude that Green's model is invalid as a criticism of Charnov's model of perfect information; when patch depression is important Charnov's marginal value theorem remains a valuable tool for understanding optimal patch exploitation.

The definition of sampling I have used may not always coincide with one's intuition about sampling. For example, sampling is often thought of as a information-gathering behaviour which sacrifices present gain to increase future gains (Houston et al. in press). The comparison I have made is (simply) between strategies which use new information and those which do not. This limitation does not affect my criticism of the patch sampling literature, because this is the traditional comparison it makes. However, this objection may limit the generality of these results.

In summary I have used my definition of the value of sampling to argue that the contemporary models of patch sampling have considered extreme cases; in natural situations where foragers are able to recognize alternative patch types, and when there is depression in patches, sampling may cause less significant departures from the perfect information non-sampling models than originally suggested by theoretical work on patch sampling (Oaten 1977, Green 1980, Iwasa et al 1981, McNamara 1982). However, I do not contend that this means that patch sampling will be unimportant. I develop a simple model of

partial patch recognition, which predicts, in a simple case, how animals should use information about patch size. I show that this is consistent with the field observations of Pyke (1982) on the foraging behaviour of bumblebees. The model is in the middle ground between Charnov's perfect patch recognition model and Oaten's perfect patch ignorance model. To my knowledge, it is the only empirical support for patch sampling of any sort; it is significant that it is only obtained by modifying Oaten's model towards Charnov's.

CHAPTER ELEVEN

Part II--Conclusions and Discussion: Ignorance is Cheap

11.1 Conclusions

In Part II I have considered several aspects of information in optimal foraging theory. I list the major conclusions and some implications of Part II here.

1. I classify the information problem into its parts, and I find that it is composed of many sub-problems. Information can enter foraging theory in many ways. Only the simplest information problems have been studied and even these may be mathematically recondite. If there seems to be no single answer to the information problem in optimal foraging, it is probably because there is no single question. In chapter seven, I outline eight information problems each of which requires a generally different type of solution.

2. I study a simple model of environmental tracking. This model may be understood in terms of a trade-off between two types of errors: sampling errors vs. over-run errors. As sampling errors become expensive with respect to over-run errors the optimal sampling frequency decreases; that is, foragers should track their environment most carefully when sampling is cheap relative to the cost of missing a change in the environment. I test this prediction experimentally with great tits (Parus major) foraging in an aviary: I find that the trend in sampling is as predicted; it changes with the ratio of cost of sampling to over-run errors in the predicted way. However, the absolute amount of sampling is too small; the birds did not track their environment as carefully as predicted. These results suggest that the predicted trade-off exists, but more realistic models are also required. Two further approaches are suggested. Firstly, the theory itself needs to

be generalized to include imperfect recognition by sampling and more complicated rules for environmental change. Secondly, these modifications will undoubtedly require esoteric mathematics; so that, more may be lost in clarity than is gained in the analysis. It is important, therefore, that simple natural systems be analysed. Problems should be sought where the mathematics can be built around the natural history. I conclude that the present model tends to over-value sampling.

3. I consider the problem of sampling patches. I criticize conventional models of patch sampling, because they have tended to consider situations in which the value of sampling is unnaturally inflated. To do this I define the value of patch sampling. I reach three conclusions using this definition. Firstly, I consider the conclusion that as contagion in the mixing distribution of patch sub-types increases that the value of sampling increases. I find that this is true, but that when contagion is great very simple and presumably cheap sampling procedures are also very effective. I call this self-moderation. Secondly, I find that if the forager can recognize certain alternative patch types that the value of sampling can be reduced by the presence of these recognizable alternatives. Thirdly, Green (1980) has considered a patch sampling model without patch depression; I show that results based on such a model inflate the value of sampling, and consequently that this model is not a fair criticism of deterministic models with patch depression. Finally, I consider simple and natural examples of patch sampling incorporating the problem of partial patch recognition. This illustrates that patch sampling will be important in some natural situations. Again, I conclude that simple models may over-value sampling.

11.2 Ignorance is cheap.

Every chapter in part II leads in some way to the conclusion that sampling and the complex modifications to foraging theory that it necessitates may be less important than previously imagined (cf. Pyke et al. 1977).

In chapter seven I show that straightforward recognition and the use of correlated information are under-studied in parts of the information problem. In chapter eight, it is seen that the models which are simplest and most understandable are those which over-value sampling. In chapter nine, I show that great tits do indeed seem to value sampling less than predicted by these simple models. In chapter ten, I examine the assumptions of patch sampling models and how they have over-valued sampling.

At a more general level, I have argued from the work of Gould, and from my definition of the value of sampling, that when information is valued in terms of the decision problem at hand, that its value is limited and finite. The intuitive view that information is of great importance to foraging animals in the way that it is important to statisticians is misleading. Animals can only value information in terms of how it increases their foraging success. Moreover, information (by sampling especially) can be expensive and difficult to obtain. On the other hand, ignorance is always cheap.

I acknowledge that this conclusion is a bit glib. All the models discussed have pre-supposed that the forager already has a ^egreat deal of information (e.g. it knows the mixing distribution of sub-types). Moreover, there are some cases where sampling is clearly significant; for example, Pyke's (1982) bumblebee data discussed in section 10.6.

I do not wish to suggest that sampling is unimportant, but that it is probably not the panacea that will make all the qualitative results found in the test of the theory (table 1.1) quantitative.

RISK VS. INFORMATION

In chapter six and the present chapter I have discussed my results in detail. However, there are two points to be made which fall neither under risk nor under information; these points deal with either general conclusions or the relationship between the two problems. I briefly discuss each point in a separate section below.

11.3 Complication.

Throughout this thesis I have approached complex natural problems by considering isolated special cases which (it is hoped) are heuristic. However, I have criticized the conventional literature of patch sampling for making unnatural and confounding simplifications. It is only fair, therefore, that I should admit the possibility that I am guilty of the same error. For example, patch depression could interact with the presence of alternatives to increase the value of sampling (although I have no reason to expect this). I acknowledge here the potential weakness of my piecemeal approach. However, as Stephens and Charnov (in press) have recently argued that the gains in clarity and analysis may outweigh this lack of generality.

One difficulty of my approach and separating risk from information, is that it may lead to neglecting possible interaction between the two problems. This is a potentially fascinating theoretical problem, which might be of importance. For example, will a risk-averter value information more than a risk-taker?, or track its environment more closely? Such questions may yield to a fairly direct application of Gould's (1974) ideas.

11.3 The importance of risk.

It is somewhat surprising that information should have entered the foraging literature before risk, because, in many ways, risk is the most direct implication of stochasticity. More importantly it is also likely to be the most generally important stochastic problem. It has already yielded more to empirical study than information has, and it is in every way the more tractable of the two problems. Moreover, the programme of work to follow up the recent burgeoning of risk is clearly marked. There are obvious experiments which need to be done (see chapter six), and obvious holes in the present theoretical framework. In short, if I were to advise workers interested in stochasticity about how to divide their time I would point them toward the problems or risk, although information is a pot worth keeping on the fire.

SUMMARY OF PARTS I AND II

Except for isolated comments I have treated risk and information in entirely separate parts. The reader may come to the end of part II having lost the thread of what was said in part I. I provide this summary as an aid to putting the elements of the problem back together.

I divide the topic of stochasticity in optimal foraging theory into two elements: risk and information. Risk is the problem of optimal choice amongst known but probabilistic outcomes. Information is the problem of how foragers should change their behaviour to gain information about ambiguous aspects of their foraging environment.

Part I - Risk

In chapter two, I review the theoretical basis of risk. Specifically, I discuss utility theory and previously suggested biological ideas about risk. In chapter three I review the empirical evidence for risk-sensitive foraging preferences. Animals show both risk-proneness (preference for risk) and risk-aversion (preference for certainty). Operant psychologists have tended to find risk-proneness, but behavioural ecologists have found risk-aversion. An important empirical result is the expected-energy budget rule: be risk-prone if your expected energy budget is negative; be risk-averse if your expected energy budget is positive.

In chapters four and five I consider risk theoretically. I show that a model based on minimization of the probability of starvation due to a short-fall predicts the expected-energy budget rule. I extend this model to consider optimal choices from complex mean-variance sets, and I use this generalization to illustrate the model's application to a conventional foraging problem. In chapter five I consider limitations

of the theory. I consider the effect of allowing starvation to occur in "mid-day" as well as at "night-fall." This change has the effect of limiting the desirability of increasing variance. I consider the importance allowing the forager to change its mind. This increases the pay-off due to the extreme-variance rule. Finally, I consider a model which allows food carried-over to the next day to have a positive value. This change tends to increase the value of risk-prone behaviour. In chapter six the limitations and implications of these models of risk are discussed.

Part II- Information

In chapter seven I review the literature on information and classify the problem and its elements. I conclude that it is a complicated and many-faceted problem.

In chapter eight I study a simple model of environmental tracking. I propose that there is a fundamental trade-off between sensitivity to environmental change and the cost of sampling. In chapter nine I test this idea using great tits foraging in an aviary. I find that the trend in sampling is as predicted, but that the birds are less sensitive to environmental change than predicted.

In chapter ten I consider the problem of patch sampling. I develop a definition of the value of sampling. I use this definition in three ways. I consider the prediction that the value of sampling increases as variance (or contagion) in the mixing distribution of patch sub-types increases. I find that this is true, but that in cases of high contagion simple sampling rules are also very effective. I consider how the presence of a simple and unambiguous patch type affects the value of sampling. I find that it can decrease the value of sampling. I also study the interaction between sampling and patch

depression. I show that patch depression can also reduce the value of sampling. I conclude that previous models have often over-valued sampling. Finally, I consider a simple and natural model of patch-use which allows partial patch recognition, this model has some empirical support.

In chapter eleven I discuss the limitations and implications of my treatment of information.

Literature Cited

Allais,M (1953) Le comportement de l'homme rationnel devant le risque: critique des postulats et axiomes de l'ecole americaine. *Econometrica* 21: 503-546.

Andersson,M. (1978) Optimal foraging area: size and allocation of search effort. *Theor. Popul. Biol.* 13: 397-409.

Andersson,M. (1981) Central place foraging in the S^xacicola rubetra. *Ecology* 52: 538-544.

Arnold,S.J.(1978) The evolution of a special class of modifiable behaviors in relation to environmental pattern. *Am Nat.* 112: 415-427.

Barnard,C.J. and Brown,C.A.J. (1981) Prey size selection and competition in the common shrew (Sorex araneus L.). *Behav. Ecol. Sociobiol.* 8: 239-243.

Barnard,C.J. and Stephens,H. (1981) Prey size selection by lapwings Vanellus vanellus in lapwing-gull Larus ridibundus associations. *Behaviour* 77: 1-22.

Belovsky,G.E.(1978) Diet optimization in a generalist herbivore, the moose. *Theor. Popul. Biol.* 14: 105-134.

Best,L. and Bierzychudek,P.(1982) Pollinator foraging on foxglove (Digitalis purpurea): a test of a new model. *Evolution* 36: 70-79.

Bobisud,L.E. and Potratz,C.J. (1976) One-trial versus multi-trial learning for a predator encountering a model mimic system. *Am. Nat.* 110: 121-128.

Bond,A.B.(1980) Optimal foraging in a uniform habitat: the search mechanism of the green lacewing Chrysopa carnea. *Anim. Behav.* 28: 10-19.

Brown, J.L. (1964) The evolution of diversity in avian territorial systems. *Wilson Bull.* 76: 160-169.

Brown, J.L. and Orians, G.H. (1970) Spacing patterns in mobile animals. *Ann. Rev. Ecol. Syst.* 1: 239-262.

Caraco, T. (1980) On foraging time allocation in a stochastic environment. *Ecology* 61: 119-128.

Caraco, T. (1981) Energy budgets, risk and foraging preferences in dark-eyed juncos (Junco hyemalis). *Behav. Ecol. Sociobiol.* 8: 213-217.

Caraco, T., Martindale, S., and Whittam, T.S. (1980) An empirical demonstration of risk-sensitive foraging preferences. *Anim. Behav.* 28: 820-830.

Carlson, A. and Moreno, J. (1981) Central place foraging in the wheatear Oenanthe oenanthe an experimental test. *J. Anim. Ecol.* 50: 917-924.

Charnov, E.L. (1976a) Optimal foraging: the marginal value theorem. *Theor. Popul. Biol.* 9: 129-136.

Charnov, E.L. (1976b) Optimal foraging: the attack strategy of a mantid. *Am. Nat.* 110: 247-259.

Charnov, E.L., Orians, G.H., and Hyatt, K. (1976) The ecological implications of resource depression. *Am. Nat.* 110: 247-259.

Coleman R. (1977) Stochastic processes. George, Allen & Unwin, London.

Conte, S. D. and de Boor, C. (1965) Elementary numerical analysis: an algorithmic approach. McGraw-Hill, New York.

Cook, R.M. and Cockrell, B.J. (1978) Predator ingestion rate and its bearing on feeding time and the theory of optimal diets. *J. Anim. Ecol.* 47: 529-548.

Corbet, S.A., Cuthill, I., Fallows, M., Harrison, T., and Hartley, G. (1981) Why do nectar foraging bees and wasps work upwards on inflorescences? *Oecologia (Berl)* 51: 79-83.

Cowie,R.J. (1977) Optimal foraging in Great Tits (Parus major).
Nature 268: 137-139.

Cowie,R.J.,and Krebs,J.R. (1979) Optimal foraging in patchy environments. In: (eds.) Anderson,R.M.,Turner,B.D. and Taylor, L.R. The British Ecological Society Symposium Vol.20 Population Dynamics:London, England. Blackwell Scientific Publications, Oxford.

Cox, D. R. and Miller H. D. (1965) The theory of stochastic processes. Chapman and Hall, London.

Davidson,D. (1978) Experimental tests of the optimal diet in two social insects. Behav. Ecol. Sociobiol. 4: 34-41.

Davies,N.B. (1977)a Prey selection and the search strategy of the spotted flycatcher Muscicapa striata: A field study of optimal foraging. Anim. Behav. 25: 1016-1033.

Davies, N.B. (1977)b Prey selection and social behaviour in wagtails (Aves: Motacillidae). J. Anim. Ecol. 46: 37-57.

Davies,N.B. and Houston,A.I. (1981) Owners and satellites: the economics of territory defence in the pied wagtail Motacilla alba J. Anim. Ecol. 50: 157-180.

DeBenedictis,P.A. ,Gill, F.B., Hainsworth, F.R., Pyke,G.H., and Wolf,L.L. (1978) Optimal meal size in hummingbirds. Am. Nat. 112: 301-316.

DeGroot,M.H. (1970) Optimal Statisical Decisions. McGraw-Hill, New York.

Dixon, W. J. and Massey, F.J. (1967). Introduction to statistical analysis. McGraw-Hill, New York.

Dixon, W. J. (1979) BMDP-79: biomedical computer program; P-series. Univ. of Calif. Press, Berkeley.

Dunstone, N. and O'Connor, R.J. (1979) Optimal foraging in an amphibious mammal Mustela vison. 1. The aqualung effect. Anim. Behav. 27: 1182-1194.

Ebersole, J.P. and Wilson, J.C. (1980) Optimal foraging: the responses of Peromyscus leucopus to experimental changes in processing time and hunger. Oecologia (Berl) 46: 80-85.

Elner, R.W. and Hughes, R.N. (1978) Energy maximization in the diet of the shore crab Carcinus maenas J. Anim. Ecol. 47: 103-116.

Erichsen, J.T., Krebs, J.R. and Houston, A.I. (1980) Optimal foraging and cryptic prey. J. Anim. Ecol. 49: 271-276.

Estabrook, G.F. and Jespersen, D.C. (1974) Strategy for a predator encountering a model-mimic system. Am. Nat. 108: 443-457.

Evans, R.M. (1982) Efficient use of food patches at different distances from the colony in blackbilled gulls. Behaviour 79: 23-78.

Feller, W. (1950) An introduction to probability theory and its applications. Vol. I. Third ed. John Wiley and Sons, New York.

Gardner, M.B. (1981) Mechanisms of size selectivity by planktivorous fish: a test of hypotheses. Ecology 62: 571-578.

Gibson, R.M. (1980) Optimal prey size selection by three-spined sticklebacks Gasterosteus aculeatus : a test of the apparent size hypothesis. Z. Tierpsychol. 52: 291-307.

Giller, P. (1980) The control of handling time and its effects on the foraging strategy of a heteropteran predator, Notonecta. J. Anim. Ecol. 49: 699-712.

Giraldeau, L.A. and Kramer, D.L. (1982) The marginal value theorem: a quantitative test using load size variation in a central place forager. Anim. Behav. in press.

Goss-Custard, J.D. (1977) Optimal foraging and the size selection of worms by redshank Tringa totanus in the field. Anim. Behav. 25: 10-29.

Goss-Custard, J.D. (1977)b Predator responses and prey mortality in the redshank Tringa totanus(L.) and a preferred prey Corophium volutator (Pallas) J. Anim. Ecol. 46: 21-36.

Gould, J.D. (1974) Risk, stochastic preference and the value of information. J. Econ. Theory 8: 64-84.

Gould, J.J. and Lewontin, R.C. (1979) The spandrels of San Marco and the Panglossian paradigm: a critique of the adaptationist programme. Proc. R. Soc. Lond. B. 205: 581-598.

Green R. F. MS Stopping rules for optimal foragers.

Green, R. F. (1980) Bayesian birds: a simple example of Oaten's stochastic model of optimal foraging. Theor. Popul. Biol. 18: 244-256.

Harmes, R. and Vickers, W. T. (MS) Optimal foraging theory and hunter-gatherer societies.

Harley, C. B. (1981) Learning the evolutionary stable strategy. J. theor. Biol. 61: 611-633.

Hartling L. K. and Plowright, R. C. (1979) Foraging by bumblebees (Bombus atratus) on patches of artificial flowers: a laboratory study. Can. J. Zool. 57(10): 1866-1870.

Hegner, R.E. (1982) Central place foraging in the white-fronted bee eater. Anim. Behav. in press.

Heinrich, B. (1976) The foraging specializations of individual bumblebees. Ecol. Monogr. 46: 105-128.

Heinrich, B. (1979a) Bumblebee Economics. Harvard University Press. Cambridge, Mass.

Heinrich, B. (1979b) Resource heterogeneity and patterns of movement in foraging bumblebees. *Oecologia (Berl)* 40: 235-245.

Heithaus, E.R. and Fleming, T.H. Foraging movements of a frugivorous bat, *Carollia perspicillata* (Phyllostomatidae). *Ecol. Monogr.* 48: 127-143.

Heller, R. and Milinski, M. (1979) Optimal foraging of sticklebacks Gasterosteus aculeatus on swarming prey. *Anim. Behav.* 27: 1127-1141.

Herrnstein, R.S. (1964) Aperiodicity as a factor in choice. *J. Exp. Anal. Behav.* 7: 178-182.

Hey, J.D. (1979) *Uncertainty in Microeconomics*. New York University Press, New York.

Hodges, C.M. (1981) Optimal foraging in bumblebees (Bombus appositus) hunting by expectation. *Anim. Behav.* 29: 1161-1171

Hodges, C.M. and Wolf, L.L. (1981) Optimal foraging in bumblebees.: Why is nectar left behind in flowers? *Behav. Ecol. Sociobiol.* 9:41-44.

Houston, A.I., Krebs, J.R., and Erichsen, J.T. (1980) Optimal prey choice and discrimination time in the great tit (Parus major). *Behav. Ecol. Sociobiol.* 6: 169-175.

Houston, A.I. and McNamara, J.M. in press The sequential approach to risk-taking. *Anim. Behav.*

Houston, A.I., McNamara, J.M. and Kacelnik, A. (in press) Some simple learning rules for acquiring information. In: McFarland, D.J. (ed.) *Functional Ontogeny*. Pitmans, London.

Howell, D.J. and Hart, D.L. (1980) Optimal foraging in Glossophagine bats: when to give up. *Am. Nat.* 115: 696-704.

Hubbard, S.F. and Cook, R.M. (1978) Optimal foraging by parasitoid wasps. *J. Anim. Ecol.*:593-604.

Hubbard, S.F., Cook, R.M., Flover, J.G. and Greenwood, J.J.D. (1982) Apostatic selection as an optimal foraging strategy. *J. Anim. Ecol.* 51 in press

Hughes, R.N. (1979) Optimal diets under the energy maximization premise: the effects of recognition time and learning. *Am. Nat.* 113: 209-222.

Hughes, R.N. and Elner, R.W. (1979) Tactics of a predator Carcinus maenas and morphological responses of the prey Nucella lapillus. *J. Anim. Ecol.* 48: 65-78.

Hughes, R.N. and Seed, R. (1981) Size selection of mussels by the Mar. *Ecol. Prog. Ser.* 6: 89-90.

Hughes, R.N. and Townsend, C.R. (1981) Maximizing net energy returns from foraging in (eds) C.R. Townsend and P. Calow *Physiological Ecology*. Blackwell Scientific Publications, Oxford.

Iwasa, Y., Higashi M. and Yanamura N. (1981) Prey distribution as a factor determining the choice of optimal foraging strategy. *Am. Nat.* 117: 710-723.

Jaeger, R.G. and Barnard, D.E. (1981) Foraging tactics of a terrestrial salamander Plethodon cinereus choice of diet in structurally simple environments. *Am. Nat.* 117: 639-664.

Janetos, A.C. and Cole, B.J. (1981) Imperfectly optimal animals. *Behav. Ecol. Sociobiol.* 9: 203-210.

Jenkins, S.H. (1980) A size distance relation in food selection by beavers Castor canadensis. *Ecology* 61: 740-746.

Kacelnik, A. (1979) Studies of foraging behaviour and time-budgeting in great tits (*Parus major*). Unpub. D.Phil. thesis, Oxford University.

Kacelnik, A., Houston, A.I. and Krebs, J.R. (1981) Optimal foraging and territorial defence in the great tit Parus major. Behav. Ecol. Sociobiol. 8: 35-40.

Kaufman, L.W. and Collier, G. (1981) Economics of seed handling. Am. Nat. 118: 46-60.

Keeney, R.L. and Raiffa, H. (1976) Decisions with multiple objectives: preferences and value trade-offs. Wiley, New York.

Killeen, P.R. (1981) Averaging theory. in (eds) C.M. Bradshaw, Szabadi, E., and C.F. Lowe. Quantification of steady state operant behaviour. Elsevier North Holland.

Killeen, P.R., Smith, J.P. and Hanson, S.J. (1981) Central place foraging in Rattus norvegicus. Anim. Behav. 29: 64-70.

Kislalioglu, M. and Gibson, R.N. (1976) Some factors governing prey selection by the fifteen-spined stickleback Spinachia spinachia. J. Exp. Mar. Biol. Ecol. 25: 159-170.

Kramer, D.L. and Nowell, W. (1980) Central place foraging in the eastern chipmunk Tamus striatus. Anim. Behav. 28: 772-778.

Krebs, J.R. (1978) Optimal foraging decision rules for predators in Krebs, J.R. and Davies, N.B. (eds) Behavioural Ecology: An Evolutionary Approach. Blackwells, Oxford.

Krebs, J.R. (1980) Optimal foraging, predation risk and territory defense. Ardea 68: 83-90.

Krebs, J.R. and Avery, M.I. in press. Test of central place foraging in a single-prey loader.

Krebs, J.R., Kacelnik, A. and Taylor, P. (1978) Test of optimal sampling by foraging great tits. Nature 275: 27-31.

Krebs, J.R., Erichsen, J.T., Webber, M.I., and Charnov, E.L. (1977) Optimal prey selection in the great tit Parus major Anim. Behav. 25: 30-38.

Krebs, J.R., Stephens, D.W. and Sutherland, W.J. in press. Perspectives in optimal foraging theory. in Clark, G.A. (ed) Centennial volume of the American Ornithologists Union.

Kushlan, J. A. (1978) Nonrigorous foraging by robbing egrets. Ecology 59: 649-653.

Larkin, S.B.C., (1981) Time and energy in decision making. D.Phil thesis, Oxford University.

Lea, S.E.G. (1979) Foraging and reinforcement schedules in the pigeon: optimal and nonoptimal aspects of choice. Anim. Behav. 27: 875-886.

Lester, N.P. in prep. The "feed-feed" decision: how goldfish solve the patch depletion problem.

Leventhal, A.M., Morrell, R.F. Morgan, E.F. and Perkins, C.M. (1959) The relation between mean reward and mean reinforcement. J. Exp. Psychol. 57: 284-287.

Lewis, A.R. (1980) Patch use by gray squirrels and optimal foraging. Ecology 61: 1371-1379.

Lobel, P.S. and Ogden, J.C. (1981) Foraging by the herbivorous parrotfish Sparisoma radians. Marine Biol. 64: 173-184.

MacArthur, R.H. and Pianka, E. R. (1966) On the optimal use of a patchy environment. Am. Nat. 100: 603-609.

Mackintosh, N.J. (1974) The psychology of animal learning. Academic Press, London.

Maynard Smith, J. (1974) Models in ecology. Cambridge University Press, Cambridge.

Maynard Smith, J. (1978) Optimization theory in evolution. Ann. Rev. Ecol. Syst. 9: 31-56.

Maynard Smith, J. and Price G.R. (1973) The logic of animals conflict. Nature 246: 15-18.

Martindale, S. (1982) Nest defence and central place foraging: a model and experiment. *Behav. Ecol. Sociobiol.* 10: 85-90.

McCleery, R.H. (1978) Optimal behaviour sequences and decision-making. In: Krebs, J.R. and Davies, N.B. (eds.) *Behavioural Ecology: an evolutionary approach*. Blackwells, Oxford.

McFarland, D.J. (1977) Decision making in animals. *Nature* 269: 15-21.

McFarland, D.J. and Houston, A.I. (1981) *Quantitative ethology: the state space approach*. Pitmans, London.

McNair, J.N. (1979) A generalized model of optimal diets. *Theor. Popul. Biol.* 15: 159-170.

McNair, J.N. in press. Energy reserves and the risk of starvation : some foraging strategies. *Am. Nat.*

McNamara, J.M. (1982) Optimal patch use in a stochastic environment. *Theor. Pop. Biol.* 21: 269-288.

McNamara, J.M. in press Control of the diffusion coefficient of a simple diffusion process in mathematics of operations research.

McNamara, J.M. MS. Optimal control of the diffusion coefficient of a simple diffusion process. (to appear as a technical report: School of Mathematics, University of Bristol.

McNamara, J. and Houston, A.I. in press. Short-term behaviour and life-time fitness. in McFarland, D.J. (ed) *Functional Ontogeny*. Pitmans, London.

Milinski, M. and Heller, R. (1978) Influence of a predator on the optimal foraging behavior of sticklebacks Gasterosteus aculeatus *Nature* 275: 642-644.

Milton, K. (1979) Factors influencing leaf choice by howler monkeys: a test of some hypotheses of food selection by generalist herbivores. *Am. Nat.* 114: 1362-368.

Mittelbach, G.G. (1981) Foraging efficiency and body size: a study of optimal diet and habitat use by bluegills. *Ecology* 62: 1370-1386.

Oaten, A. (1977) Optimal foraging in patches: a case for stochasticity. *Theor. Popul. Biol.* 85: 673-690.

Ollason, J.G. (1980) Learning to forage optimally? *Theor. Popul. Biol.* 18: 44-56.

Orians, G.H. and Pearson, N.E. (1979) On the theory of central place foraging. in (eds) Horn, D.J., Mitchell, R. and Stair, G.R. *Analyses of Ecological Systems*. Ohio State University Press.

Oster, G.F. and Wilson, F.O. (1978) *Caste and ecology in the social insects*. Princeton Univ. Press. Princeton, New Jersey.

Palmer, A.R. (1979) Prey selection by non-visual invertebrate predators: a field test of optimal foraging. Ph.D. thesis, University of Washington.

Parker, G.A. (1978) Searching for mates. In: (eds.) Krebs, J.R. and Davies, N.B. *Behavioural Ecology: An Evolutionary Approach*. Blackwell Scientific Publications, Oxford.

Pastorok, R.A. (1980) The effects of predator hunger and food abundance on prey selection by Chaoborus larvae. *Limnol. Oceanogr.* 25: 910-921.

Pearson, N.E. (1974) Optimal foraging theory. *Quan. Science Paper No. 39*, Center for Quan. Science in forestry, fisheries and wildlife. University of Washington. Seattle, Washington.

Pleasants, J.M. (1981) Bumblebee response to variation in nectar availability. *Ecology* 62: 1648-1661.

Pubols, B.H. (1962) Constant versus variable delay of reinforcement. *J. Comp. Physiol. Psychol.* 55: 52-56.

Pulliam, H.R. (1974) On the theory of optimal diets. *Am. Nat.* 108: 59-75.

Pulliam, H.R. (1980) Do chipping sparrows forage optimally? *Ardea* 68: 75-82.

Pulliam, H.R. and Dunford, C. (1980) Programmed to learn. Columbia University Press, N.Y.

Pulliam, H.R. and Millikan, G. (in press) Avian social organization in the non-reproductive season. In: Farner, D.S. and King, J.R. (eds.) *Avian Biology* (vi) Academic Press, New York.

Pyke, G.H. (1978a) Optimal foraging in bumblebees and co-evolution with their plants. *Oecologia (Berl)* 36: 281-294.

Pyke, G.H. (1978b) Optimal foraging in hummingbirds: testing the marginal value theorem. *Am. Zool.* 18: 739-752.

Pyke, G.H. (1979) The economics of territoriality in the golden-winged sunbird. *Am. Nat.* 114: 131-145.

Pyke, G.H. (1981) Why hummingbirds hover and honeyeaters perch. *Anim. Behav.* 29: 861-877.

Pyke, G.H. (1982) Foraging in bumblebees: rules of departure from an inflorescence. *Can. J. Zool.* 60: 417-428.

Pyke, G.H., Pulliam, H.R. and Charnov, E.L. (1977) Optimal foraging: a selective review of theory and tests. *Q. Rev. Biol.* 52: 137-154.

Raiffa, H. (1968) *Decision Analysis*. Addison-Wesley, Mass.

Rapport, D.J. (1980) Optimal foraging for complementary resources. *Am. Nat.* 116: 324-346.

Real, L.A. (1980a) Fitness uncertainty and the law of diminishing returns in evolution and behaviour. *Am. Nat.* 115: 623-638.

Real, L.A. (1980b) On uncertainty and the law of diminishing returns in evolution and behavior. In: Staddon, J.E.R. (ed.) *Limits to action: the allocation of individual behavior*. Academic Press, New York.

Ringler, N.H. (1979) Selective predation by drift feeding brown trout Salmo trutta. J. Fish. Res. Board Can. 36: 392-403.

Schoener, T.W. (1971) Theory of feeding strategies. Ann. Rev. Ecol. Syst. 2: 379-404.

Sibly, R.M. and McFarland, D.J. (1976) On the fitness of behavior sequences. Am. Nat. 110: 607-617.

Sih, A. (1980) Optimal behavior: can foragers balance two conflicting demands? Science 210: 1041-1043.

Sih, A. (1980) Optimal foraging: partial consumption of prey. Am. Nat. 116: 281-290.

Staddon, J.E.R. (1980) Optimality analyses of operant behaviour and thier reaction to optimal foraging. in (ed) Staddon, J.E.R. Limits to Action: The Allocation of Individual Behaviour. Academic Press.

Stein, R.A. (1977) Selective predation, optimal foraging, and the predator-prey interaction between fish and crayfish. Ecology 58: 1237-1253.

Stephens, D.W. (1981) The logic of risk-sensitive foraging preferences. Anim. Behav. 29: 628-629.

Stephens, D.W. and Charnov, E.L. (1982/in press) Optimal foraging: some simple stochastic models. Behav. Ecol. Sociobiol.

Tinbergen, J. (1981) Foraging decisions in starlings (Sturnus vulgaris L.) Ardea 69: 1-67.

Townsend, C.R. and Hildrew, A.G. (1980) Foraging in a patch environment by a predatory net spinning caddis Plectrocnemia conspersa larva: a test of optimal foraging theory. Oecologia (Berl) 47: 219-221.

Vadas, R.L. (1977) Preferential feeding: an optimization strategy in sea urchins. Ecol. Monog. 47: 337-371.

Von Neumann, J. and Morgenstern, O. (1947) Theory of games and economic behavior. Second edition. Princeton University Press, Princeton.

Waddington, K.D., Allen, T., and Heinrich, B. (1981) Floral preference of bumblebees (Bombus edwardsii) in relation to intermittent versus continuous rewards. Anim. Behav. 29: 779-784.

Waddington, K.D. and Heinrich, B. (1979) The foraging movements of bumblebees on vertical "inflorescences": an experimental analysis. J. Comp. Physio. 134: 113-117.

Waddington, K.D. and Holden, L.R. (1979) Optimal foraging on flower selection by bees. Am. Nat. 114: 179-196.

Werner, E.E. and Mittelbach, G.G. (1981) Optimal foraging: field tests of diet choice and habitat switching. Am. Zool. 21: 813-830.

Werner, E.E. and Hall, D.J. (1974) Optimal foraging and the size selection of prey by the bluegill sunfish Lepomis macrochirus. Ecology 55: 1042-1052.

Whitham, T.G. (1977) Coevolution of foraging Bombus- nectar dispensing Chilopsis: a last dreg theory. Science 197: 325-326.

Williams, D. (1982) Studies of optimal foraging using operant techniques. Ph.D. thesis, Univ. of Liverpool.

Winterhalder, B.P. and Smith E.A. (1981) (eds) Hunter-gatherer foraging strategies. University of Chicago Press, Chicago.

Zach, R. (1979) Shell dropping, decision making and optimal foraging in northwestern crows. Behaviour 68: 106-117.

Zimmerman, M. (1979) Optimal foraging, a case for random movement. Oecologia (Berl) 43: 261-267.

Zimmerman, M. (1981) Optimal foraging, plant density and the marginal value theorem. Oecologia (Berl) 49: 148-153.

Appendix I: Risk-Aversion in the Great Tit.

D.W. Stephens and R.C. Ydenberg.

This appendix records the materials and methods used in the risk-sensitivity experiment of Stephens and Ydenberg. The purpose of the experiment was to study sequential aspects of risk-sensitivity. This aspect of the experiment failed. However, the results do show a strong tendency for risk-aversion in the great tit.

A.I. Materials and Methods

Subjects The experimental subjects were six great tits (Parus major), three males and three females. The birds were kept in separate metre-cube cages. The cages were kept in a small room, which has a controlled light:dark cycle 10:14. The birds were fed a diet of commercial food (Aleckwa Mischung I) with added boiled egg and grated carrot. The birds were food deprived after dark each evening to insure that they were hungry for the next days experiment.

The experiment involved complicated rules to determine the amount of deprivation each bird had before its test; these rules are not relevant to the data presented. The experiments were conducted between 1030 hrs. and 1500 hrs. Food was always returned to the birds between 15 and 30 minutes after finishing their daily experiment.

Experimental Apparatus The experimental chamber was a modified metre-cube cage. The plan of this cage is shown in figure AI-1. One wall of the cage was replaced by a plywood and smoked perspex wall. The wall had holes through which drawers (trays) slid. An observer could present food by putting fly pupae in a petri dish held by an indentation of the tray and sliding the tray into the cage. The petri dishes were covered by small opaque discs to prevent the bird from seeing the contents of the petri dish before deciding which drawer to feed from. The birds quickly learned to knock these discs away. The birds were taught that food would only be presented when they perched on a central perch which was situated equidistant from both trays.

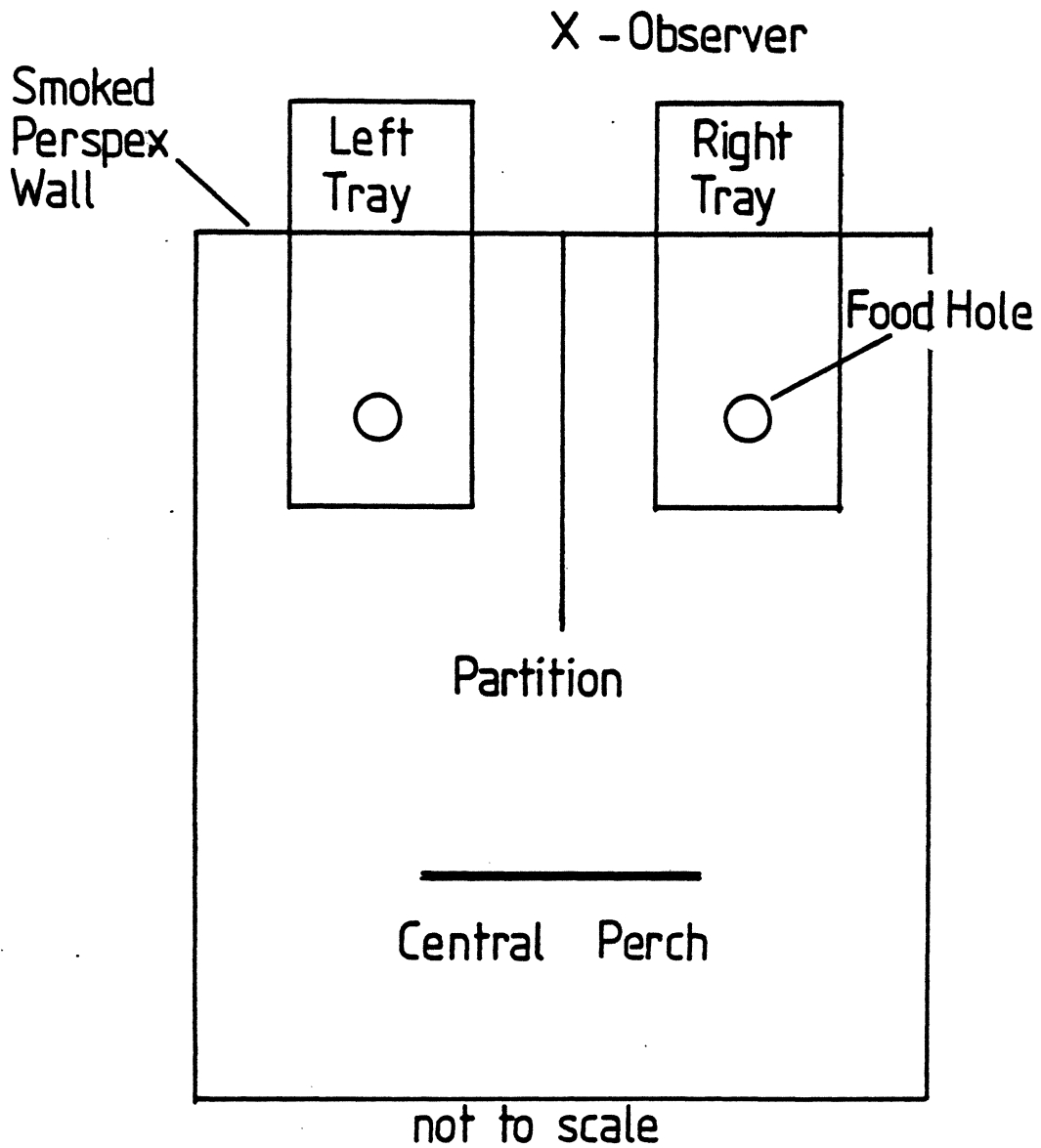


Figure AI.1 Plan of Experimental Cage

The experiment The experiment consisted of four steps: (1) The birds were taught that both trays provided 2 fly pupae per trial by sliding the trays in one at a time in a randomly determined order until each bird sampled ten trials on the left side and ten trials on the right side. (2) A preference test was conducted by presenting both trays simultaneously. The bird's choices were recorded for 20 such simultaneous presentations. (3) The birds were then taught that the tray that they preferred in step (2) provided seven fly pupae one-half of the time and one pupa the other half; and the previously unpreferred tray always provided four pupae. Again this was taught by ten trials for each drawer presented singly and in random order. (4) The birds were then given another preference test as discussed in step (2). However, the number of trials varied (due to the considerations of the experimental design). The results are shown in table 3.2.

- Purnell, D. J. & Thompson, J. N. Jr. 1973. Selection for asymmetrical bias in a behavioural character of *Drosophila melanogaster*. *Heredity*, 31, 401-405.
- Reeve, E. C. R. 1960. Some genetic tests on asymmetry of sternopleural chaeta number in *Drosophila*. *Genet. Res., Camb.*, 1, 151-172.

(Received 10 November 1980; MS. number: sc-90)

The Logic of Risk-Sensitive Foraging Preferences

It has recently been shown that the foraging preferences of animals are sensitive to variance in the probability distribution of a food reward (Caraco et al. 1980). This is an important result for optimal foraging models, because they often assume that the expected value of the food reward distribution is the quantity that should be maximized and that variance may be safely ignored (see Pyke et al. 1977 for a review). Caraco (1980) has suggested that some aspects of risk sensitivity may be explained by minimization of the probability of starvation, where starvation is defined as falling short of some threshold food requirement at the end of the day. In this note I formalize this alternative optimality hypothesis and show analytically that it accounts for the existence and some aspects of the nature of the risk sensitivity reported by Caraco et al. (1980).

Caraco et al. (1980) offered yellow-eyed juncos (*Junco phaeonotus*) a choice between two feeding stations. One station always provided a constant number of seeds, while the alternative station provided a variable number determined by a probability distribution. The number of seeds at the constant station was equal to the expected number provided at the variable station. Caraco et al. (1980) demonstrated that juncos preferred the constant station (or were risk-averse) when their expected daily energy intake exceeded their expected daily energy expenditure. More surprisingly, they also demonstrated that juncos preferred the variable station (or were risk-prone) when their expected daily energy expenditure exceeded their expected daily intake. Juncos show preferences which are consistent with a simple rule: be risk-averse if your expected 24-h energy budget is positive, be risk-prone if your expected 24-h energy budget is negative. This rule has been suggested, on intuitive grounds, by Caraco (1980). In this note I show that the expected energy budget rule follows directly from minimization of the probability of starvation.

Consider a small bird foraging in winter. Imagine that at least R units of food reserve are required to survive the night. Let S_0 be the random variable that describes the bird's food reserve when it becomes too dark to forage. The probability of starvation is equivalent to $P(S_0 < R)$. (This assumes that starvation only occurs at night, which may not be true; however the results are not affected as long as food reserves are unlikely to become dangerously low during the day.) Assume that the day's foraging is divided into n intervals. At the beginning of the i th interval the forager makes some decision about how to forage and receives X_i units of food. The X_i 's are random variables and are assumed to be independently distributed with $E(X_i) = \mu$ and $\text{Var}(X_i) = \sigma^2$. Assume that exactly r units of food are expended during each interval (this assumes that r is independent of the foraging decision, which cannot be strictly true). Let S_n be the bird's internal food supply when there are n decisions left before night-fall. We can express the reserves at the end of the day as

$$S_n = S_0 + \sum_{i=1}^n X_i - nr.$$

The probability of starvation is

$$P(S_0 < R) = P\left(\sum_{i=1}^n X_i < R - S_n + nr\right). \quad (2)$$

The central limit theorem for independent random variables states that for n sufficiently large, the random variable

$$Z = \frac{\sum_{i=1}^n X_i - \sum_{i=1}^n \mu}{\sqrt{\sum_{i=1}^n \sigma_i^2}}$$

has a normal distribution with a mean of zero and a variance of one. For convenience we define the following

variables: the mean expected value $\bar{\mu} = \frac{1}{n} \sum_{i=1}^n \mu_i$,

the mean variance

$$\bar{\sigma}^2 = \frac{1}{n} \sum_{i=1}^n \sigma_i^2, \text{ and } \bar{\sigma} = \sqrt{\bar{\sigma}^2}.$$

Also let

$$t = \frac{1}{\bar{\sigma}} \left[\frac{(R - S_n) + n(\bar{\mu} - \bar{\mu})}{\sqrt{n}} \right].$$

It follows from the central limit theorem and appropriate substitutions into equation (2) that

$P(\text{starving}) = P(S_0 < R) = P(Z < t) = \Phi(t)$, (3) where $\Phi(t)$ is the cumulative distribution function (cdf) of the standard normal distribution, $N(0, 1)$. By definition of the cdf of an absolutely continuous random variable, $\Phi(t)$ increases in a strictly monotonic way with t . This means that as t increases, the probability of starvation increases. Therefore, we can restrict our attention to effects on t .

Consider the partial derivative of t with respect to $\bar{\sigma}$,

$$\frac{\partial t}{\partial \bar{\sigma}} = \frac{1}{\bar{\sigma}^2} \left[\frac{(S_n - R) + n(\bar{\mu} - r)}{\sqrt{n}} \right]. \quad (4)$$

The sign of $\partial t / \partial \bar{\sigma}$ is the same as the sign of the term $(S_n - R) + n(\bar{\mu} - r)$.

Therefore

$$\partial t / \partial \bar{\sigma} > 0 \text{ where } \bar{\mu} - r > (R - S_n)/n, \quad (5a)$$

$$\partial t / \partial \bar{\sigma} < 0 \text{ where } \bar{\mu} - r < (R - S_n)/n. \quad (5b)$$

These are the asymptotic (large n) conditions for risk sensitivity. Expression (5a) is the condition for risk aversion, because the probability of starvation increases with increasing variance. Expression (5b) is the condition for risk proneness, because the probability of starvation decreases with increasing variance. This is an analytical derivation of the 24-h expected energy budget rule of Caraco (1980). This result also shows the following. (a) Indifference to risk is unlikely, because it should only occur in the unlikely situation where expression (5) becomes an exact equality. (b) The direction of the change in the probability of starvation with increasing variance is independent of the magnitude of the variance. Figure 1

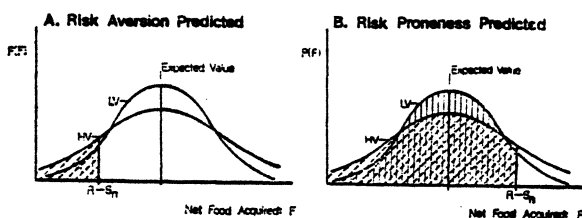


Fig. 1. The figure shows two normal probability distributions of the net amount of food acquired during the day. They have equal expected values but different variances. The area under the curves to the left of $R - S_n$ is the probability of starvation, because $R - S_n$ is the amount of food required to survive. The area marked with sloping dashed lines represents the probability of starvation if high variance (HV) is chosen; vertical solid lines are for the low variance (LV) choice. The rule derived in the text is simply the observation that: (Case A) when the expectation is greater than the requirement, the probability of starvation is smaller for low variance; and (Case B) when the expectation is less than the requirement, the probability of starvation is less for high variance.

presents a graphical argument which illustrates this derivation.

Expression (5) may be unnecessarily strict. If n is large, $\lim (R - S_n)/n = 0$, and these conditions become:

$n \rightarrow \infty$

$$\begin{aligned} &\text{be risk-averse if } \bar{\mu} > r, \\ &\text{and be risk-prone if } r > \bar{\mu}. \end{aligned} \quad (6)$$

When many decisions are left to make, the forager may base its risk-sensitivity on a simpler rule of thumb. Expression (6) states that if environmental conditions are such that the forager's reserves are expected to increase, the forager should be risk-averse. If reserves are expected to decrease, the forager should be risk-prone.

The 24-h expected energy budget rule holds because the total energy acquired, $T = \sum_{i=1}^n X_i$, is normally distributed. T will probably be normal, because it will be the sum of many independent random variables and should approach normality according to the central limit theorem. If n is not large enough, or if T is non-normal for other reasons, the rule may not hold. In cases where the mean and variance are closely related, the 24-h expected energy budget rule will be too simplistic.

A forager's risk-sensitivity may also be influenced by the number of decisions left to make, n . Expression (5) suggests that a forager with only a few decisions left will become risk-prone at a higher reserve level than will a forager with many decisions left to make. Risk sensitivity should have a sequential component. This is a characteristic of many procedures in statistical decision theory (DeGroot 1970).

The results presented show that an hypothesis based on the minimization of the probability of starvation explains both risk-averse and risk-prone foraging preferences. It is not generally true that 'a risk-prone forager accepts a greater probability of starvation in attempting to acquire a greater net energetic intake' (Caraco 1980). The model shows that there are situations, i.e. (5b), where the forager may decrease its probability of starvation by risk-prone behaviour.

I am grateful to Tom Caraco for helpful discussions, stimulating manuscripts, and lunch. A. I. Houston, J. R. Krebs, R. C. Ydenberg and the referees provided needed criticism. E. L. Charnov pointed out the importance of mean/variance relationships. A manuscript by J. McNamara and A. I. Houston has helped shape my ideas.

D. W. STEPHENS

Edward Grey Institute,
Department of Zoology,
South Parks Road
Oxford OX1 3PS.

References

- Caraco, T. 1980. On foraging time allocation in a stochastic environment. *Ecology*, 61, 119-128.
- Caraco, T., Martindale, S. & Whitham, T. S. 1980. An empirical demonstration of risk-sensitive foraging preferences. *Anim. Behav.*, 28, 820-830.
- DeGroot, M. H. 1970. *Optimal Statistical Decisions*. New York: McGraw-Hill.
- Pyke, G. H., Pulliam, H. R. & Charnov, E. L. 1977. Optimal foraging: a selective review of theory and tests. *Q. Rev. Biol.*, 52, 137-154.

(Received 3 September 1980; revised 19 November 1980;
MS. number: sc-82)

Giving-Up as a Poisson Process: The Departure Decision of the Green Lacewing

Predators that forage for aggregated prey appear to require a decision rule for determining the point at which to discontinue their search in a given prey patch and move on to another. Although the optimum rule depends heavily on features of the searching behaviour of the predator and the distribution of the prey (Oaten 1977), most previous authors have assumed that the decision must involve an assessment of the capture rate within a patch and a comparison with the mean capture rate in the environment as a whole (Krebs 1978). When the perceived quality of the given patch becomes significantly less than the expected quality of the next one, the predator should leave. Because the time interval since the last prey capture is the most readily available measure of the instantaneous capture rate, it has been suggested that foraging animals may monitor this interval and leave the patch when it exceeds some critical value (Krebs 1978). The 'giving-up time', by this argument, should be uniform across patches within a habitat and inversely proportional, across habitats, to the mean prey availability. Although this inference has been supported by empirical studies, Cowie & Krebs (1979) have recently suggested that the correlation could be a sampling artifact. Even if departure from a patch were independent of the interval between prey encounters, the mean giving-up time would still be shorter, on the average, in a rich environment than in a poor one. A re-analysis of several experiments on patch foraging by predatory insects, described in detail elsewhere (Bond 1980), can be used to test Cowie & Krebs' independence hypothesis.

The predatory larva of the green lacewing, *Chrysopa carnea* Stephens, responds to an encounter with either a prey item or a branch terminus with a transient increase in searching intensity and thoroughness. Under the assumption that an animal's hunger level is its simplest accessible indicator of mean prey abundance, differences in prey availability were simulated by

Optimal Foraging: Some Simple Stochastic Models

D.W. Stephens¹ and Eric L. Charnov²

¹ Edward Grey Institute, Department of Zoology, University, South Parks Road, Oxford OX1 3PS, England

² Department of Biology, University of Utah, Salt Lake City, Utah 84112, USA

Received November 9, 1981 / Accepted April 24, 1982

Summary. Some simple stochastic models of optimal foraging are considered. Firstly, mathematical renewal theory is used to make a general model of the combined processes of search, encounter, capture and handling. In the case where patches or prey items are encountered according to a Poisson process the limiting probability distribution of energy gain is found. This distribution is found to be normal and its mean and variance are specified. This result supports the use of Holling's disc equation to specify the rate of energy intake in foraging models. Secondly, a model based on minimization of the probability of death due to an energetic shortfall is presented. The model gives a graphical solution to the problem of optimal choices when mean and variance are related. Thirdly, a worked example using these results is presented. This example suggests that there may be natural relationships between mean and variance which make solutions to the problems of 'energy maximization' and 'minimization of the probability of starvation' similar. Finally, current trends in stochastic modeling of foraging behavior are critically discussed.

However, stochastic models have an important place in optimal foraging theory for three reasons.

1. Many biologically important phenomena are fundamentally stochastic, so that deterministic models do not make sense (e.g. sampling).

2. Stochastic models can be compared to deterministic models to tell us how good, as approximations, these simpler models are. This comparison is useful both for models of inherently stochastic problems as well as for models that can reasonably be modeled deterministically.

3. Conversion of deterministic to stochastic models by replacing deterministic variables with expectations of random variables may be inappropriate because characteristics of the random distribution other than the expectation may be critical; for example, both expectation and variance of food intake may be important.

The most valuable stochastic models may be useful in all three (or other) ways. To make this clear we give examples of these uses of stochastic models below.

Introduction

Foraging animals search for, encounter, and handle their prey according to rules which are most accurately described by statements of probability. That is, we expect stochastic models of foraging behavior to resemble nature more closely than deterministic models. However, many models of optimal foraging behavior are deterministic. There are two reasons for this. Deterministic models are nearly always simpler than their stochastic counterparts, and it is often easy to convert deterministic into stochastic models by representing distributions of random variables by their expectations (or mean values).

Inherently Stochastic Problems

Some foraging problems are essentially stochastic and cannot be modeled deterministically. For example, deterministic models usually assume that the forager 'knows,' in some sense, the relevant characteristics of its environment. This might be reasonable in a relatively static environment, such as those usually provided in experimental tests of foraging theory (cf. Krebs et al. 1977). However, in a stochastic world, parameters such as encounter rate may change unpredictably. Some time and energy will have to be invested in sampling the environment because real environments will often have unpredictable and hence stochastic characteristics. The trade-

Ms. No. 328 Author Stephens Ms. 45 Pages 1-13

Springer-Verlag, Heidelberg / H. Stürtz AG, Würzburg

Provisorische Seitenzahlen Provisional page numbers I. Korr.

Date 4.6.82

Trv

off between the information gain and the energy gain due to a foraging decision can only be analyzed stochastically.

Comparison and Analysis

Simple deterministic models may be reasonable approximations to more realistic stochastic models. One use of stochastic models, and perhaps the most valuable one, is to explore the extent to which deterministic approximations are reasonable. Excellent examples of this are the re-analyses of Oaten's (1977) 'information by patch exploitation' model by Green (1980) and McNamara (in press). Green and McNamara demonstrated not only that the optimal patch residence for Oaten's naive (stochastic) forager may differ from Charnov's (1976a) knowledgeable (deterministic) forager, but they have also shown which factors influence the degree of this difference, for example, variance in the distribution of patch quality.

In this paper we present a stochastic argument which allows us to compare conventional deterministic models with a stochastic case. We use ideas from the mathematical theory of renewal (Cox 1962) to make a generalized model of the combined processes of search, encounter, capture and handling. We use these results to derive expressions for the mean and variance of energy gain due to foraging for a fixed time. We find the limiting distribution of energy gain due to foraging.

Stochasticizing by Expectations

Recent empirical (Caraco et al. 1980; Real 1981; Caraco 1981; Waddington et al. 1981) and theoretical work (Caraco 1980; McNamara and Houston, in press; Real 1980a, b; Stephens 1981) suggests that conversion of deterministic models by simply considering deterministic variables to be the means of probability distributions may be inappropriate: foragers are, and should be, sensitive to both the mean and variance of the distribution of food reward. This has been called risk-sensitive foraging by Caraco et al. (1980). Following the terminology of economics (e.g., Keeney and Raiffa 1976) Caraco et al. called preference for variance in food reward risk-prone-ness and preference for certainty risk-aversion.

Both risk-prone (Leventhal et al. 1959; Pubols 1962; Herrnstein 1964) and risk-averse (Real 1981; Waddington et al. 1981) preferences have been demonstrated in foraging animals. It may surprise biologists that both risk-prone and risk-averse preferences occur. Real (1981a, b), for example, suggests that risk-aversion ought to be universal. The key

evidence is provided by Caraco et al. (1980) and Caraco (1981). Caraco et al. were able to manipulate the energy budgets of yellow-eyed juncos (*Junco phaeotus*), and offered them a constant and a variable food reward, where the expected value of the variable reward was equal to the value of the constant reward. The juncos preferred the constant choice when their expected energy budgets were positive, but preferred the variable choice when their expected energy budgets were negative. Caraco et al. demonstrated the following simple empirical rule: Be risk-averse if your expected daily energy budget is positive; be risk-prone if your expected daily energy budget is negative.

Caraco (1980) suggested that risk-sensitive foraging behavior would be predicted by a model minimizing the probability of starvation. Stephens (1981) has shown analytically that a simple model minimizing the probability of falling short of some threshold food requirement predicts the expected energy budget rule. In these cases there are compelling reasons to abandon simple conversion by expectations and develop full scale stochastic models which allow foragers to choose between variances as well as means.

We present a graphical generalization of the Stephens (1981) model of risk-sensitive foraging preferences. We consider how a forager minimizing its probability of starvation might best choose from a relatively complex set of mean-variance pairs.

Results

Our results consist of three models. Firstly, we attempt to establish a fairly general framework for comparing conventional deterministic models to a stochastic case. To do this we use results from renewal theory to find the limiting (i.e., time large) distribution of energy gain. Secondly, we consider a simple case in which elements of this distribution of energy gain, other than its mean, may be important: that is, when natural selection minimizes the probability of death due to a short-fall. Finally, we sketch an example which applies both of these results to a conventional foraging problem. The choice of an optimal patch residence time. To aid continuity brief discussion and summary paragraphs follow each model.

The Limiting Distribution of Energy Gain

Earlier papers (Charnov 1976a, b) used deterministic arguments to derive equations for the rate of energy intake by a predator attacking individual prey items or traveling between patches or clumps of prey. In

this section we formulate the problem as a stochastic process through time (Charnov 1973; McNair 1979). The random variable of interest is the net energy intake in some time interval, 0 to τ . As general results, the mean and variance of this variable, may be obtained, at least for large τ . Since the basic argument is similar for patches or prey types, we shall develop it only for prey types.

The Model. Our basic technique is to consider the time from leaving one prey item until leaving the next as a random event. Suppose that the predator began hunting at time zero and has just finished handling a prey item. Let this be the $r-1^{\text{th}}$ item handled. Let T_r be the time elapsed between leaving the $r-1^{\text{th}}$ item and leaving the r^{th} item. T_r is the sum of the search time for the r^{th} item and its handling time. The beginning of the search for the next item can be called a renewal point or event. Renewal theory (Cox 1962) is the study of the properties of this type of process. The T_r are drawn from the same probability distribution and are assumed to be independent of one another. This is a repeating environment assumption (MacArthur 1972).

To make this more precise, we shall use the following definitions: h_r = handling time for the r^{th} prey item. T_r = the search time between leaving the $r-1^{\text{th}}$ and encountering the r^{th} prey items (thus $T_r = h_r + T_{sr}$). $E(X)$ = expectation of an arbitrary random variable X . $V(X)$ = variance of an arbitrary random variable X . N_τ = number of prey items encountered and handled in a time interval of length τ . ξ_r = the net energy gain associated with prey item r . The ξ_r are identically distributed and independent with finite mean and variance. In general, the ξ_r are not independent of the T_r .

$$G_\tau = \sum_{r=1}^{N_\tau} \xi_r, \text{ the sum of the } \xi_r \text{ for } N_\tau$$

prey encountered and handled during τ . We shall not count in G_τ any prey the predator has encountered, but has not yet finished handling as this prey is not counted in N_τ . G_τ is the net energy intake during τ .

Because they will be used in several final results we define the following simplified notation:

$$\begin{aligned} E(G_\tau) &= \mu_\tau & E(T_r) &= \mu & E(h_r) &= \mu_h & E(\xi_r) &= \mu_\xi \\ V(G_\tau) &= \sigma_\tau^2 & V(T_r) &= \sigma^2 & V(h) &= \sigma_h^2 & V(\xi) &= \sigma_\xi^2 \\ \text{Cov}(T_r, \xi_r) &= \sigma_{r\tau} \end{aligned}$$

G_τ can be considered as a cumulative renewal process and Smith's theorems (Smith 1955) may be used to obtain the asymptotic (τ goes to infinity) mean and variance. The basic results may be found in Cox (1962, p. 100). We simply quote them here:

$$\mu_\tau = E(G_\tau)_{\tau \rightarrow \infty} \approx \frac{\tau \mu_\xi}{\mu} \quad (1.1)$$

$$\sigma_\tau^2 = V(G_\tau)_{\tau \rightarrow \infty} \approx \tau \left[\frac{\sigma_\xi^2}{\mu} + \frac{\sigma^2 \mu_\xi^2}{\mu^3} - 2 \frac{\mu_\xi \sigma_{r\tau}}{\mu^2} \right]. \quad (1.2)$$

Paloheimo (1967, 1971) was the first to use these results to model predator behavior, although he did not consider the accumulated catch $[E(G_\tau)]$ in terms of energy intake. Even though these are derived under the assumption that τ is large, the approximations may be good for 'small' intervals, depending on how fast the stochastic process converges. G_τ can also be shown to follow a normal distribution as a result of the central limit theorem. To see how these results may be used to generate energy intake equations, consider the case where prey items are encountered 'at random.'

Let there be k prey types, with type i having associated with it two random variables. The first is the handling time (h_i), the second is the energy value (e_i) associated with the handling time events. These have been previously discussed (Charnov 1973, 1976b). Their means are h_i^* and e_i^* . Each type forms an encounter process with the predator. The encounter processes are assumed to be independent Poisson processes through time with encounter rates λ_i ($i=1, 2, \dots, k$). During the handling of a prey item, no new item may be encountered. It is important to notice that it is the encounter process which is Poisson, that is we expect the number of encounters in a fixed search time to fit a Poisson distribution not the number of prey captured in a fixed time foraging.

If we look at the encounter process formed by pooling the k processes (i.e., look at an encounter regardless of type), the resulting pooled process is still Poisson and the new rate parameter is $\lambda^* = \sum_{i=1}^k \lambda_i$. The probability that an encounter is with a prey type j is λ_j/λ^* and is independent of search time (T_{sr}) (Cox 1962).

With the above in mind, the search time between encounters (T_{sr}) has a negative exponential distribution with mean and variance:

$$E(T_{sr}) = 1/\lambda^* \quad V(T_{sr}) = 1/\lambda^{*2}. \quad (1.3)$$

We have for the r^{th} renewal

$$E(h_r) = \mu_h = [\sum \lambda_i h_i^*] / \lambda^* \quad (1.4)$$

$$E(\xi_r) = \mu_e = [\sum \lambda_i e_i^*] / \lambda^* \quad (1.5)$$

$$V(\xi_r) = \sigma_e^2 = [\sum \lambda_i E(e_i^2)] / \lambda^* - \mu_e^2 \quad (1.6)$$

$$V(h_r) = \sigma_h^2 = [\sum \lambda_i E(h_i^2)] / \lambda^* - \mu_h^2 \quad (1.7)$$

$$\begin{aligned} Cov(T_r, \xi_r) = \sigma_{re} &= Cov(h_r + T_{sr}, \xi_r) \\ &= Cov(h_r, \xi_r). \end{aligned} \quad (1.8)$$

Now:

$$T_r = T_{sr} + h_r, \text{ so } \mu = E[T_{sr}] + \mu_h$$

or

$$\mu = 1/\lambda^* + [\sum \lambda_i h_i^*] / \lambda^*. \quad (1.9)$$

Since

$$\begin{aligned} Cov(T_{sr}, \xi_r) &= 0, \quad \sigma^2 = V[T_{sr}] + \sigma_h^2 \\ \sigma^2 &= 1/\lambda^{*2} + \sigma_h^2. \end{aligned} \quad (1.10)$$

If we substitute (1.5) and (1.9) into (1.1) the following results as $\tau \rightarrow \infty$.

$$\frac{\mu_r}{\tau} \approx \frac{\sum \lambda_i e_i^*}{1 + \sum \lambda_i h_i^*} \quad (1.11)$$

substituting (1.5) and (1.10) into (1.2) gives the variance for Z_r as $\tau \rightarrow \infty$.

$$\sigma_z^2 \approx \tau \left[\frac{\sigma_e^2}{D} + \frac{\mu_e^2(1 + \sigma_h^2 \lambda^{*2}) \lambda^*}{D^3} - \frac{2\mu_e \sigma_{re}}{D^2} \right]$$

$$\text{where } D = 1 + \sum \lambda_i h_i^*. \quad (1.12)$$

Expression (1.11) and (1.12) completely specify the limiting distribution of energy gain, G_r , because Smith's theorems (Smith 1955) show that the limiting distribution of a cumulative renewal process approaches normality according to the central limit theorem.

Implications of the Model. The chief importance of this result is to specify the distribution of energy gain. However, this result is also important as further stochastic justification of expression (1.11). Expression (1.11) is the form of Holling's disc equation (Holling 1959) most commonly used in models of optimal foraging. Many authors have used expression (1.11) as a currency of maximization in deterministic models of optimal diet choice (Charnov 1976b; Maynard Smith 1974; McNair 1979; Pearson 1974; Pulliam 1974; Schoener 1969; Werner and Hall 1975) and models of optimal patch use (Charnov 1976a; Orians and Pearson 1979; Pyke et al. 1977 for a review of this and other currencies).

when encounters are assumed to be at random. Previously expression (1.11) has been justified for stochastic situations because it is equal to the expected gain per prey encounter, $E(\xi_r)$, over the expected time spent per encounter (including interpret time), $E(T_r)$. The average amount of energy gained per time unit approaches $E(\xi_r)/E(T_r)$ as the number of encounters becomes infinite, according to the law of large numbers (Gilliam et al., in press; Turelli et al., in press; McNair, personal communication). The result presented here adds the weight of the central-limit theorem to the use of this formulation for foraging models. Usually, the central-limit theorem implies faster convergence to the limiting state than the law of large numbers. However, the comparison is not perfect because the renewal justification is asymptotic in time, while the law of large numbers justification is asymptotic in the number of prey encountered. Important unanswered questions are: How fast is the asymptotic distribution reached (but see Turelli et al., in press) and what affects the speed of approach?

Conclusions. In this section we find the limiting probability distribution of energy gain for a foraging animal when encounters with prey items (or patches) occur at random. This distribution is normal and we give its mean and variance. The results also support the use of Holling's disc equation to represent the rate of energy intake in foraging models.

The Risk of Starvation and Optimal Preferences

Many species have nonbreeding seasons in their life histories; during these periods survival until the next breeding season may be the most important attribute of fitness. Stephens (1981) has shown that when death due to starvation is defined as falling short of some threshold food requirement, both the mean and variance of the food reward distribution are important in determining optimal behavior. However, Stephens's formulation only considers the case where the mean is fixed and the forager is allowed to choose from a set of different variances. In this section we present a graphical model which extends this formulation by specifying sets of mean-variance pairs which provide equal probabilities of starvation.

The Model. To illustrate the model we consider a small bird foraging in winter. Many small birds suffer significant overnight mortality in winter (e.g., Prys-Jones 1979) and most must stop feeding at nightfall. Assume that the bird's main problem is

surviving from day to day, by acquiring sufficient energy reserves to survive the night.

Let R be the amount of food required to survive the night, and let S_0 be the random variable which describes the bird's energy supply when it must stop foraging for the night. The probability of starvation is $P(S_0 < R)$. Assume that the forager makes decisions about how to forage at intervals. Let k be the number of decision intervals left in the day. Let S_k be the bird's reserves when there are k foraging decisions left in the day. Let X_i be the net amount of food gained in the i^{th} interval, where N_i is a random variable with distribution

$$f(n_i | \mu_i, \sigma_i^2) \mu_i = E(X_i) \quad \text{and} \quad \sigma_i^2 = V(X_i).$$

Assume that the X_i are independently distributed. Then

$$S_0 = S_k + \sum_{i=1}^k X_i \quad (2.1)$$

therefore

$$P(S_0 < R) = P\left(S_k + \sum_{i=1}^k X_i < R\right) = P\left(\sum_{i=1}^k X_i < R_{\text{net}}\right)$$

where R_{net} is the net food requirement, ($R_{\text{net}} = R - S_k$). Now if k is large we can use the central limit theorem for independent random variables to find the distribution of the sum of random variables $\sum_{i=1}^k X_i$. For convenience, we define: $\mu_\tau = \sum_{i=1}^k \mu_i$, the total expected gain due to the decisions made in the remaining time, τ ; $\sigma_\tau = \left[\sum_{i=1}^k \sigma_i^2\right]^{1/2}$, the standard deviation of the distribution of the remaining food reward. The central-limit theorem states that

$$Z = \frac{\sum_{i=1}^k X_i - \mu_\tau}{\sigma_\tau} \quad (2.2)$$

has an approximately standard normal distribution, with mean zero and variance one. We use Z algebraically to show that $\mu_\tau + \sigma_\tau Z$ is probabilistically

equivalent to $\sum_{i=1}^k X_i$, that is

$$\begin{aligned} &= P(\text{Starvation}) = P\left(\sum_{i=1}^k X_i < R\right) \\ &= P(\mu_\tau + \sigma_\tau Z < R_{\text{net}}) \\ &= P[Z < (R_{\text{net}} - \mu_\tau)/\sigma_\tau] \\ &= \Phi[(R_{\text{net}} - \mu_\tau)/\sigma_\tau]. \end{aligned} \quad (2.3)$$

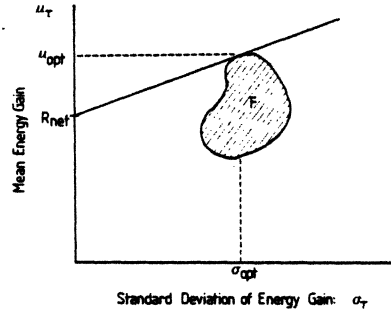


Fig. 1. The general graphical solution for optimal preferences from complex sets of mean and variance. The hatched shape, marked F , represents an arbitrary feasible set of mean-standard deviation pairs. The pair which minimizes the probability of death due to an energetic short-fall is found by constructing the line originating at R_{net} (the net energy requirement) on the mean axis and which passes through, or touches, the feasible set with the greatest slope possible

Where $\Phi(\cdot)$ is the cumulative distribution functions (cdf) of a standard normal. This means that the probability of starvation strictly increases with the term $(R_{\text{net}} - \mu_\tau)/\sigma_\tau$.

Let $Z' = (R_{\text{net}} - \mu_\tau)/\sigma_\tau$. Any fixed value of Z' defines a set of pairs (μ_τ, σ_τ) ; all members of such a set give the same probability of starvation. A forager minimizing its risk of death due to starvation should be indifferent to a choice between any two points lying on the line

$$\mu_\tau = R_{\text{net}} - \sigma_\tau Z'. \quad (2.4)$$

When Z' is fixed, (2.4) is the familiar slope-intercept equation of a straight line in $\mu_\tau - \sigma_\tau$ space. This suggests the simple graphical method of solution illustrated in Fig. 1. In general, the forager will be able to choose from a limited feasible set of mean-variance pairs, which is illustrated by the hatched blob in the figure. Consider the family of lines, each of which intersects the μ_τ axis at R_{net} . The line of greatest slope which can be drawn through the feasible set gives the smallest possible probability of starvation, and the point, or points, (μ_τ, σ_τ) of the feasible set through which this line passes is (are) the optimal mean-standard deviation pair (or pairs). The line of greatest slope gives the solution, since when the slope, $-Z$, is largest the value Z , which we want to minimize, is smallest. In a particular optimality problem we would have calculated the feasible set by considering the effects of varying some behavioral parameter of interest, e.g., patch residence time, or the probability of pursuit upon encounter. These calculations would allow us to work back-

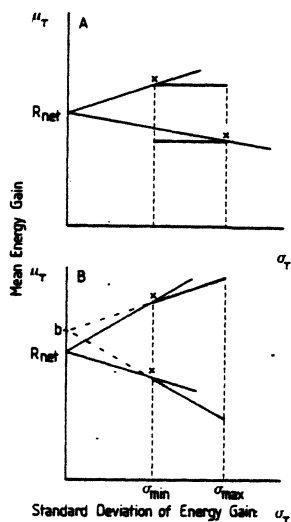


Fig. 2A, B. Linear Feasible Sets. A Two horizontal feasible sets, i.e., sets in which the forager chooses from a range of variances (standard deviations in the graph) at a constant mean. This illustrates that the extreme variance rule, discussed in the text, is a special case of this graphical model. Crosses indicate the optimal choices. B Example of the general rules for linear feasible sets, i.e., when $\mu_r = m\sigma_r + b$, $\sigma_{min} \leq \sigma_r \leq \sigma_{max}$. The optimal preference depends only on the position of the intercept b relative to the net requirement, R_{net} . Note that in the lowest feasible set the extreme variance rule does not hold because all possible means are less than the net requirement, R_{net} , but the optimal policy is still to minimize variance (and consequently to maximize the mean)

ward from the optimal (μ_r, σ_r) pair to find the optimal behavioral (or choice) parameter. This procedure is illustrated by a worked example in the next section.

The simple behavioral rule for minimizing the probability of starvation proved by Stephens (1981) can be derived as a special case of this model. Stephens showed that at a given mean a forager would minimize its probability of starvation due to a short fall by: (A) Always minimizing variance if its expected gain is greater than its requirement; (B) always maximizing variance if its requirement is greater than its expected gain. We call this the extreme variance rule. It follows from the general graphical solution because the feasible set is a horizontal line (i.e., μ_r is constant). Figure 2A shows that when $\mu_r > R_{net}$ the line of greatest slope passes through the smallest possible σ_r in the feasible set (condition A above). But when $\mu_r < R_{net}$ the line of greatest slope passes through the largest possible σ_r in the feasible set.

If the feasible set is linear, but not in general, horizontal, there is a simple extension of this rule.

Assume that the feasible set is defined by the linear equation $\mu_r = m\sigma_r + b$, where $\sigma_{min} \leq \sigma_r \leq \sigma_{max}$. It is easy to show that the probability of starvation due to short fall is minimized by: choosing σ_{max} if $R_{net} > b$, and choosing σ_{min} if $R_{net} < b$. The direction of preference for variance is independent of the slope of the feasible set. Consequently, the optimal mean value may be either the maximum possible mean or the minimum possible mean in the two cases above (the case where $R_{net} < b$ is illustrated in Fig. 2B). We have shown that the assumption of a fixed expected reward is critical to the extreme variance rule. In the lower feasible set in Fig. 2B, all possible means are less than the minimum requirement R_{net} , but the optimal choice is maximum mean and minimum variance.

Linearity is only the simplest kind of nontrivial relationship which might exist between μ_r and σ_r . This general linear solution is an advance over the special case of fixed expected gain, because it makes the point that the nature and degree of a forager's preferences over variance are changed when the forager has a restricted choice of mean-variance combinations.

Implications of the Model. Caraco (1980) suggested that a forager should be risk-prone if it has 'no available allocation strategy' which produces an expected net reward greater than expected requirements. However, Caraco used the term 'risk-proneness' to mean that a forager would prefer to draw a realization from a probability distribution than to be given the mean value of the same distribution with certainty. This usage is in line with economics and mathematics (cf. Keeney and Raiffa 1976). That is, the definitions of risk-proneness and risk-aversion have built into them the assumption that a forager is offered a choice between variances at the same mean. As we have argued above, and as the results of the first section suggest, foragers are unlikely to be offered such choices in nature. A useful theory of risk-sensitivity will need to go beyond these strict definitions of risk-proneness and risk-aversion to consider which distribution should be preferred from an arbitrary set of distributions. The model we have presented does this for an animal minimizing its probability of death due to a short fall.

In this model we begin by relating energy intake to survival. The assumption that there is some fixed threshold requirement is only the simplest possible relationship. There may, for example, be unpredictable elements in the threshold (e.g., if the overnight temperature is unpredictable, in our small bird example); or the probability of survival may be an increasing sigmoid function of daily food reward

(McNamara and Houston, in press). The results will be more complicated if the distribution of daily food reward is not normal. Departures from normality may be relatively common, especially in the tails of the approximated distribution. Perhaps most dubiously, we assume that all amounts of energy gained greater than the requirement are equivalent in terms of fitness; that is, we tacitly assume that there is no advantage to a large energetic 'carry-over.' McNamara and Houston (in press) present an interesting discussion of this problem.

These assumptions may not always make sense. The model may fit the case of an over-wintering bird, foraging on an especially cold day; but it will not fit the same bird in the breeding season. Real (1980a, b) and Caraco (1980) have presented more general models of risk-sensitivity. However, the relationship between the natural history of foragers and these general models is not always clear. Despite the limitations of our model, it is useful as a simple and analytical alternative to mean maximization models.

Conclusions. In this section we present a simple model of risk-sensitive foraging preferences. The model is based on minimization of the probability of energetic short-fall. Our model suggests that preferences between distributions of food reward ought to be sensitive to relationships between mean and variance. Our graphical model provides a technique for finding optimal preferences from complex mean-variance sets.

Patch Use and the Risk of Starvation - A Worked Example

An important question raised in the last section is what mean-variance relationships are most common and most general? In this section we apply the results of the two previous sections to finding the optimal patch residence time for an animal minimizing its risk of starvation. This illustrates how the relationship between mean and variance may be found in a foraging problem.

The Model. Consider an animal foraging in discrete patches in a manner which minimizes its probability of starvation. Assume that time spent in a patch completely specifies the net amount of food acquired from the patch. Patches are identical, and numerous enough that environmental depletion due to the forager's behavior is not significant, i.e., the forager is always entering undepleted patches. The only stochasticity is due to random variation in the time required to find a new patch.

Our problem is to find a patch residence time, t , which minimizes the probability of starvation, as discussed above. Let T_i be the random variable which represents the time between leaving one patch and entering another. Assume that T_i is exponentially distributed with parameter q , that is patches are encountered at random. Let t be the residence time, which is the parameter over which we wish to maximize, and let $g(t)$ be the negatively accelerated gain function (a deterministic function) which relates the net amount of energy gained from a patch to the residence time. We use Eq. (1.1) and (1.2) to give specific expressions for the asymptotic mean and standard deviation of energy gain. This example is simple, because for a fixed residence time, the variance in energy gain within a patch is zero, as is the covariance of the interrenewal interval and the energy gain per patch. The required terms are:

$$\mu_e = g(t)$$

$$\mu = E(T_i) + t = 1/q + t$$

$$\sigma^2 = V(T_i) = 1/q^2$$

$$\sigma_e^2 = 0$$

$$\sigma_{re} = 0.$$

By substitution we find the required mean and variance terms:

$$E(G_i | t) = \mu_i(t) = \frac{g(t)\tau}{(1/q) + t} \quad (3.1)$$

$$V(G_i | t) = \sigma_i^2(t) = \frac{\tau(1/q^2)[g^2(t)]}{[(1/q) + t]^3} \quad (3.2)$$

It is a standard result of patch-use theory that μ_i is maximized where

$$g'(t) = \frac{q(t)}{(1/q) + t} \quad (3.3)$$

This result is a special case of Charnov's (1976a, b) marginal value theorem. Consider the derivative of the standard deviation, σ_i , with residence time.

$$\hat{c}\sigma_i/\hat{c}t = (\tau^{1/2}) \cdot \left\{ \frac{g(t)[(1/q) + t]^{3/2} - (3/2)g(t)[(1/q) + t]^{1/2}}{q[(1/q) + t]^3} \right\}.$$

Since the term in the denominator is always positive the sign of the derivative is determined by the difference in the numerator. The standard deviation increases with t as long as

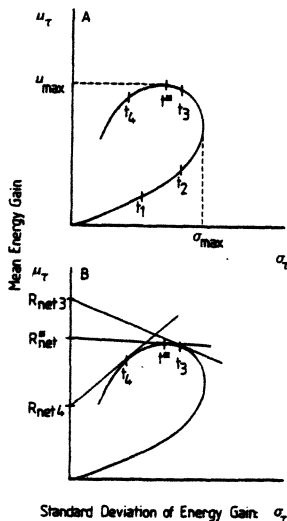


Fig. 3A, B. Patch residence time and the risk of starvation. The curve shown in A shows the shape of the feasible set for an animal foraging in discrete patches, and choosing an optimal patch residence time as discussed in the text. The patch residence times (t_i 's) marked on the curve indicate how increasing residence time affects the mean and standard deviation, that is $t_4 > t^* > t_3 > t_2 > t_1$. B Application of the graphical model of the previous section to finding optimal patch residence times. R_{net}^* is the net requirement at which mean maximization is preserved. But, as with R_{net4} , if the requirement is less than R_{net}^* , it is best to stay longer (i.e., $t_4 > t^*$); and as with R_{net3} , if the requirement is greater than R_{net}^* it is best to leave sooner (i.e., $t^* > t_3$)

$$g'(t) > \frac{(3/2)g(t)}{(1/q)+t} \quad (3.4)$$

the standard deviation decreases with t where this inequality is reversed, and is therefore at a maximum when t is such that

$$g'(t) = \frac{(3/2)g(t)}{(1/q)+t} \quad (3.5)$$

Notice the similarity between (3.3) and (3.5). We can conclude, because $g(t)$ is negatively accelerated, that the residence time at which the maximum standard deviation occurs is always less than the residence time at which the mean intake is maximized. We summarize the results of this analysis: (i) μ_r and σ_r are both unimodal functions of patch residence time, and (ii) that the maximum standard deviation will always occur at a smaller residence time than the maximum possible mean, and (iii) both μ_r and σ_r are zero when $t=0$. These points are enough to allow us to plot the feasible set in μ_r - σ_r space and use our earlier graphical treatment (Sect. 2) to make pre-

dictions about the qualitative differences between a forager choosing its patch residence time to maximize its intake rate, and a forager minimizing its probability of a shortfall.

Figure 3A shows the general shape of the feasible set. Figure 3B illustrates the qualitative deductions which can be made considering the general shape of the feasible set, and the graphical method of solution presented in Sect. 2. The optimal residence time for a forager minimizing its probability of starvation will have the following properties: (1) It will usually be near the residence time predicted by mean-maximization (the marginal value theorem) because of the feasible set's oblong shape. (2) There will be a critical level of requirement R_{net}^* at which the optimal residence time is exactly the same as in the mean-maximization case. (3) If the actual requirement is less than R_{net}^* then the optimal behavior is to stay longer than predicted by mean-maximization. (4) If the actual requirement is greater than R_{net}^* , the optimal residence time is shorter than predicted by mean-maximization. (5) As the requirement becomes larger the optimal residence time becomes shorter and approaches the time at which the maximum standard deviation is attained, but will never becomes smaller than this.

As a general rule we may say that if the requirement is low compared to expectations, then the forager should stay longer than expected by mean-maximization; if the requirement is high, the forager should stay for a shorter period than expected by mean-maximization, but never shorter than the time producing maximal variance. These predictions are only qualitative. There are indications from numerical examples that in many cases departure from the mean-maximization hypothesis may be small. The extreme variance rule does not hold in this example because neither the highest nor the lowest possible variance should be chosen, except in the limiting case. It does give us insight into why this result occurs. The forager chooses increasingly higher variance as requirements become large relative to expected gains. However, since the feasible set is not linear, the forager should never choose extremes of variance because it would have to accept a disproportionate loss in mean to do so.

Several authors have presented modifications of the marginal value theorem. Oaten (1977) and McNamara (in press) consider the case in which the forager also gains information while foraging. McNair (in press) considers a modification where the forager is allowed to encounter new patches while searching a patch. In both of these cases the optimal residence time is longer than predicted by the marginal value theorem (although the comparison with

Oaten's model requires a redefinition of the Charnov's marginal capture rate, cf. McNamara, in press). McNair (in press) also analyzes an optimal residence time model which minimizes the probability of starvation over an infinite time period. His model differs from ours in that he envisions no critical reckoning time (nightfall in our example). McNair's 'starvation' is simply running out of fuel during the foraging process. McNair finds that the probability of starvation is minimized by staying longer than the marginal value theorem would predict. McNair's infinite time model is similar to ours when net requirements are very small relative to expected gains. Our model suggests that over-staying may be the most common case; because requirements will normally be much less than expected gains. However, in the case of severe energetic stress relatively short stays may be optimal.

Conclusions. We consider the optimal patch residence time for a forager minimizing its risk of starvation. This requires that we specify the relationship between residence time, mean energy gain, and the standard deviation of energy gain. The surprising shape of the feasible set we find in this section suggests that it may be misleading to try to find general relationships between mean and variance. This is because we need to know how mean and variance are specified by a choice of foraging strategy (e.g., patch residence time). Our result suggests that this relationship may depend critically on the nature of the foraging problem at hand.

Discussion

The problem of stochasticity in optimal foraging is actually two problems. Firstly, in a stochastic environment a forager's behavior will not yield, for example, ten calories with certainty; instead a forager will at best be able to choose a probability distribution of calories. Choosing from a known set of probability distributions and choosing certain alternatives are clearly different problems. The first stochastic problem is how should foragers change their preferences when they are faced with known types of stochastic variability. Economists call this the problem of pure risk (Hey 1979).

The second stochastic problem is more complicated and less direct. Imagine that the state of nature is represented by one of two possible overlapping probability distributions, A and B. Assume that the forager knows what the alternatives are, but cannot distinguish between them before deciding what tactic to use. Further, imagine that if the forager knew it was faced with distribution A it should

use a different tactic than if faced with B. The tactic actually chosen should then reflect information (about the relative likelihoods of A vs B) gained while foraging. That is, the forager has a sampling problem, or in the terminology of economics a problem in pure uncertainty (Hey 1979), although it would probably be less confusing to call this a problem of information.

The separation of the stochastic problem into these two elements is logically important, not because animals face pure information or pure risk problems in nature, but because it suggests a sensible way for our study of stochasticity to proceed. In this paper we have discussed problems of pure risk. In more conventional terms we have made perfect information models. In this discussion we discuss risk and information and the controversies surrounding them, in turn. Finally, we consider the problem of rates in a stochastic environment.

Risk and Utility

The most general solution to the problem of risk is given by utility theory. In outline the procedure is simple. We have a utility function (exactly how we come to have it is an issue we discuss later), defined over all possible outcomes. The utility function specifies which outcomes are preferred and how much a given outcome is preferred over another. The fundamental result of utility theory is that a utility function *can* be defined so that distributions can be placed in order of preference by their expected utilities. The idea that such a function exists is called the expected utility hypothesis (cf. Raiffa 1968).

To help make the distinctions clear consider the form of a deterministic foraging problem. The steps in a deterministic analysis are: (1) choose a currency and (2) specify cost-benefit functions. (Schoener 1971). Expected utility is the same as Schoener's currency because it is the quantity we maximize. But in a risk problem cost-benefit functions do not relate behavior to expected utility. They relate behavior to the distribution of some intermediate reward (e.g., amount of food). The expectation of the utility function, given this distribution, provides the final link to the currency.

A utility analysis of foraging behavior would take the following form. (1) Choose a utility, for example, the probability of survival. All the considerations which apply to the choice of a deterministic currency apply to this decision, e.g., when is the probability of survival the most important fitness attribute? (2) Specify the relationship between the proposed utility and food. (3) Specify the relation-