An Introduction to Integration and Probability Density Functions*

Kevin Mitchell
Department of Mathematics and Computer Science
Hobart and William Smith Colleges
Geneva, NY USA 14456

March 1999

This and other related materials are available on-line at:
`http://science.holycross.edu/departments/biology/kprestwi/behavior/ESS/ESS_index_frmset.html`

*This is a supplement to Kenneth Prestwich's material "The Mixed ESS Solution to the War of Attrition."

1 Finding Expected Lifetime Net Benefits

Recall the situation under discussion at "The Mixed ESS Solution to the War of Attrition" website. Benefits are only obtained by the focal strategist when she wins—i.e., when the focal strategist is willing to pay a higher cost than her opponent from the **mix** ($x < m$, where m is the cost the focal strategist will pay).

$$\text{Net Benefit to } \mathbf{fix}(x = m) \text{ in a win} = V - x, \tag{1}$$

where V is the resource value and x is the cost the opponent from **mix** is willing to pay.

Unfortunately, (1) is not sufficient for our needs. The complexity of the war of attrition intervenes! Recall that the mix is composed of an infinite number of component strategies. $\mathbf{fix}(x = m)$ only faces one of these supporting strategies in any given contest. Thus, (1) only describes the net gain in one specific contest. You should realize that this particular contest will probably be quite rare given the many different strategists that $\mathbf{fix}(x = m)$ could face from the mix. Thus, one particular contest and its benefits will have little if any important *lifetime effect* on $\mathbf{fix}(x = m)$'s fitness. Single contests cannot describe the net benefit that the focal supporting strategy expects to gain from a large number (a lifetime) of contests.

To get an accurate measurement of lifetime net gains, we need to take into account all types (costs) of contests that $\mathbf{fix}(x = m)$ will win (i.e., those where the opposing strategy x is at most m) and the probability of each.

$$\text{Net Benefit} = \sum_{0 \leq x \leq m} [(V - x) \times (\text{Probability of facing } x)]. \tag{2}$$

Note that this is an infinite sum because there are an infinite number of different costs (strategies) x between 0 and m. To handle this type of sum, we will need to use calculus, which we now briefly consider. We will return the question of lifetime net benefits once we have introduced the appropriate calculus techniques.

2 An Introduction to Integration

The probability of facing a particular strategy x is determined by a **probability density function** or **pdf**. We will illustrate the idea by examining a number of different situations.

Situation 1:

Suppose that there are an infinite number of strategies x that players adopt with $0 < x \leq 1$. Assume further that each strategy x with $0 < x \leq 0.5$ is just as likely to occur as any other in this interval. Finally, assume that each strategy x with $0.5 < x \leq 1$ is *twice* as likely to occur as any strategy x with $0 < x \leq 0.5$. If we assign a positive probability p to any strategy x in $(0, 0.5]$, then since there are an infinite number of other equally likely startegies, they, too, would have probability p. But then summing these probabilities would produce an infinitely large value, not 1. For this reason, the probability that any particular strategy x is encountered must be 0.

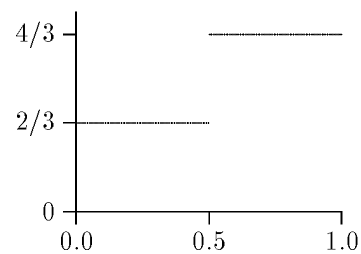
A better question to ask is what the probability is of facing a strategy x that falls within a certain interval. For example, let p denote the probability of encountering a strategy x that lies in the interval $(0, 0.5]$. I claim that $p = 1/3$. Here's why:

- The strategies in $(0.5, 1]$ are twice as likely to be encountered as those in $(0, 0.5]$, so they have a total probability $2p$, i.e., twice the probability of those in $(0, 0.5]$.

- But since the only strategies are those between 0 and 1, we must have $1 = p + 2p = 3p$, or $p = 1/3$.

Suppose, next, that we wanted to know the probablity of encountering a strategy x with $0 < x \leq 0.25$? Well, since all the strategies in $(0, 0.5]$ are equally like, and since their total probability is $p = 1/3$, and since $(0, 0.25]$ is exactly half the interval $(0, 0.5]$, we conclude that the probablity that $0 < x \leq 0.25$ must be half of p, that is $p/2 = 1/6$. In the same way you should be able to show that the probablity that $0.75 < x \leq 1$ is $1/3$.

Figure 1: A geometric realization or probability density function $p(x)$ for situation 1.



The function $p(x)$ in **Figure 1** gives a geometrical representation of situation 1. It has the following properties.

- $p(x_1) = p(x_2)$ for any two strategies x_1 and x_2 in $(0, 0.5]$ because they are equally likely.
- $p(x_1) = p(x_2)$ for any two strategies x_1 and x_2 in $(0.5, 1]$ because these, too, are equally likely.
- But $p(x_2) = 2p(x_1)$ for any strategy x_2 in $(0.5, 1]$ and x_1 in $(0, 0.5]$ because strategy x_2 is *twice* as likely as x_1.
- The region under the graph of $p(x)$ on the interval from 0 to 0.5 is rectangle whose area is $\frac{1}{2} \cdot \frac{2}{3} = \frac{1}{3}$ Which is the probablity that x is in $(0, 0.5]$.
- The region under the graph of $p(x)$ on the interval from 0.5 to 1 is rectangle whose area is $\frac{1}{2} \cdot \frac{4}{3} = \frac{2}{3}$ which is the probablity that x is in $(0.5, 1]$.
- The region under the graph of $p(x)$ over the entire interval from 0 to 1 is $\frac{1}{3} + \frac{2}{3} = 1$ which is the probablity that x is in $(0, 1]$.
- More generally, the area under $p(x)$ on the interval $[a, b]$ represents the probability of a strategy x where $a \leq x \leq b$.

The function $p(x)$ is an example of a **probability density function** or pdf. Such functions must satisfy two conditions:

1. $p(x) \geq 0$ for all x,
2. the total area under $p(x)$ must be 1.

For a given x the function $p(x)$ measures the *comparitive* or *relative likelihood* of strategy x. This is why the graph of p is twice as high on $(0.5, 1]$ as it is on $(0, 0.5]$.

Siutation 2:

Assume that the only strategies x are those such that $0 \leq x \leq 2$ and that all of these strategies are equally likely. Since they are all equally likely, $p(x)$ must be constant on $[0, 2]$, so that the region under the graph is a rectangle. Since the area under the pdf over $[0, 2]$ must be 1, the the constant height (for the rectangle) must be $\frac{1}{2}$ because the base is 2 (see **Figure 2 (a)**).

Figure 2: (a) The probability density function $p(x)$ when all strategies between 0 and 2 are equally likely. The area under the graph from a to b is $\frac{1}{2}(b-a)$.

Since all strategies are equally likely, the probability that a particular strategy x lies in $[0, 1]$ is just $\frac{1}{2}$ or half the total probablity. Of course this corresponds exactly to the area under $p(x)$ in **Figure 2 (a)** over the interval $[0, 1]$: it is a $1 \times \frac{1}{2}$ rectangle. In fact, if b is any number in $[0, 2]$, then the probability of encountering a strategy x in the interval $[0, b]$ is the area under the curve from 0 to b, that is, $b \times \frac{1}{2} = \frac{b}{2}$. The probablity that one encounters a strategy with waiting time less than or equal to x is called the the **cumulative probability distribution of quitting times**, and is denoted by $P(x)$. Note the uppercase P for the cumulative function and the lowercase p for the density function. In this example, $P(x)$ is just the area of the rectangle from 0 to x with height $\frac{1}{2}$, so

$$P(x) = x/2, \qquad 0 \leq x \leq 2.$$

If we want to find the probability of encountering a strategy x between a and b, we could find the area of the rectangle from a to b. Its base would be $b-a$ and the height $\frac{1}{2}$, so its area is $\frac{b-a}{2}$. But this probablity can be expressed more generally using $P(x)$. The probability we want is just the area from 0 to b minus the area from 0 to a as in **Figure 2 (b)**. But this is just

$$P(b) - P(a) = \frac{b}{2} - \frac{a}{2} = \frac{b-a}{2}.$$

Situation 3:

Suppose again that the only possible strategies x are those such that $0 \leq x \leq 2$ and that the pdf is $p(x) = x/2$. Notice that this is a legitimate pdf since $p(x) \geq 0$ and the region under the graph of p is a triangle whose area $\frac{1}{2} \cdot 2 \cdot 1 = 1$ (see **Figure 3 (a)**).

What is the cumulative distribution $P(x)$ for this situation? Well, $P(x)$ is just the area under p from 0 to x which is just a triangle with base x and height $p(x)$ (see **Figure 3 (b)**). Thus,

$$P(x) = \frac{1}{2} \cdot x \cdot p(x) = \frac{1}{2} \cdot x \cdot \frac{x}{2} = \frac{x^2}{4}.$$

For example, the probability that a strategy x lies in the interval $[0, 0.5]$ is

$$P(0.5) = \frac{(0.5)^2}{4} = \frac{0.25}{4} = 0.0625,$$

Figure 3: (a) The probability density function $p(x) = x/2$ for $0 \leq x \leq 2$. (b) $P(x) = x^2/4$ is the area under p from 0 to x.

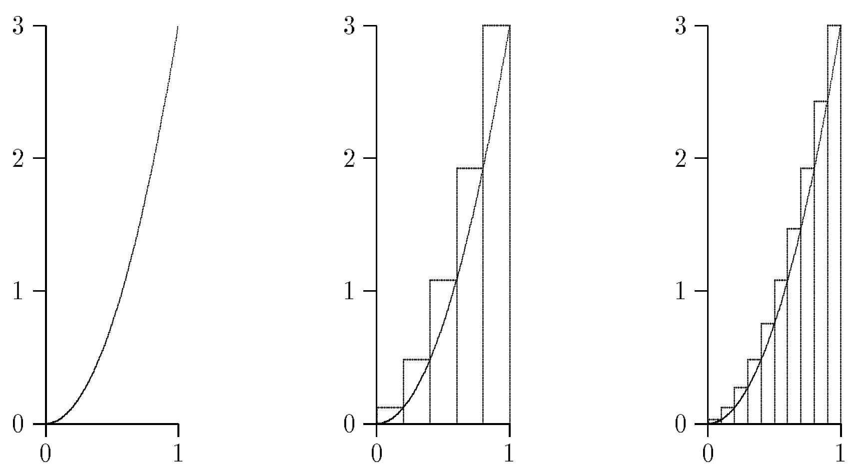
while the probability that x is in the interval $[0.5, 1.5]$ is

$$P(1.5) - P(0.5) = \frac{(1.5)^2}{4} - \frac{(0.5)^2}{4} = \frac{2.25}{4} - \frac{0.25}{4} = 0.5.$$

Situation 4:

Suppose now that the possible strategies are restricted to $0 \leq x \leq 1$ and that $p(x) = 3x^2$ (see **Figure 4 (a)**). Clearly $p(x) \geq 0$, so to show that $p(x)$ is a pdf, we need to show that the area under this curve is 1. But how do we find the area of a curved region?

Figure 4: (a) The probability density function $p(x) = 3x^2$ for $0 \leq x \leq 1$. (b) $P(x)$ is the area under p from 0 to x.



We can't directly use the area formula of a rectangle to determine the area, but we can use it indirectly to *approximate* the area under the curve. Suppose we divide the interval $[0, 1]$ into five subintervals of equal width

$$\Delta x = \frac{1-0}{5} = 0.2.$$

Approximate the area under the curve on each subinterval by using a rectangle whose height is determined by evaluating p at the right-hand endpoint of the subinterval (see **Figure 4 (b)**). These five right-hand endpoints will be $x_1 = 0.2$, $x_2 = 0.4$, $x_3 = 0.6$, $x_4 = 0.8$, and $x_5 = 1$. So in this case, the five heights would be $p(x_1) = p(0.2)$, $p(x_2) = p(0.4)$, $p(x_3) = p(0.6)$, $p(x_4) = p(0.8)$, and $p(x_5) = p(1)$. Thus, approximate area is

$$\begin{aligned}
\sum_{k=1}^{5} p(x_k)\Delta x &= p(0.2)\cdot 0.2 + p(0.4)\cdot 0.2 + p(0.6)\cdot 0.2 + p(0.8)\cdot 0.2 + p(1)\cdot 0.2 \\
&= 0.2[p(0.2) + p(0.4) + p(0.6) + p(0.8) + p(1)] \\
&= 0.2[3(0.2)^2 + 3(0.4)^2 + 3(0.6)^2 + 3(0.8)^2 + 3(1)^2] \\
&= 0.2[0.12 + 0.48 + 1.08 + 1.92 + 3] \\
&= 0.2(6.6) \\
&= 1.32.
\end{aligned}$$

If we used ten rectangles instead each with width $\Delta x = 0.1$ (see **Figure 4 (c)**), the approximation is even better. This time, the area is

$$\begin{aligned}
\sum_{k=1}^{10} p(x_k)\Delta x &= p(0.1)\cdot 0.1 + p(0.2)\cdot 0.1 + \cdots + p(0.9)\cdot 0.1 + p(1)\cdot 0.1 \\
&= 0.1[p(0.1) + p(0.2) + \cdots + p(0.9) + p(1)] \\
&= 0.1[3(0.1)^2 + 3(0.2)^2 + \cdots + 3(0.9)^2 + 3(1)^2] \\
&= 0.1(11.55) \\
&= 1.155.
\end{aligned}$$

The same process using one-hundred rectangles (no drawing for this!) yields an approximation of 1.01505 and with one-thousand rectangles the approximation is 1.0015005. It appears that these approximations are getting close to 1 as the number of rectangles gets large.

Mathematicians define the *exact* area under the curve to be the limit of this rectangular approximation proccess as the number of rectangles n becomes infinitely large,

$$\lim_{n\to\infty} \sum_{k=1}^{n} p(x_k)\Delta x.$$

We denote the limit of this summation process more compactly as

$$\int_0^1 p(x)\,dx.$$

This is read as, "The integral of $p(x)$ from 0 to 1." The integral sign, $\int$, is an elongated S, a reminder to us that an integral is really a sum. The lower and upper limits of integration (here 0 and 1, respectively) are the beginning and ending points of the interval where the sum is taking place. The expression $p(x)dx$ is meant to remind us of $p(x)\Delta x$, the area of a rectangle of height $p(x)$ and width Δx. Think of $p(x)dx$ as being the area of infinitesimally thin rectangle of height $p(x)$.

In our particular case, with $p(x) = 3x^2$, it appears that

$$\int_0^1 3x^2\,dx = 1,$$

and this is, in fact, correct. The Fundamental Theorem of Calculus tells us that under certain circumstances such integrals can be easily evaluated using functions called antiderivatives. In fact, the cumulative distribution function is just an antiderivative of $p(x)$. In this particular situtation using calculus, $P(b)$, that is, the area under $p(x) = 3x^2$ over the interval from 0 to b, is given by the formula

$$P(b) = \int_0^b 3x^2\,dx = b^3.$$

Since $P(b) = b^3$, then the cumulative distribution function is $P(x) = x^3$, which calculus students will recognize as an antiderivative of $p(x) = 3x^2$. Using methods developed in integral calculus, such antiderivatives can be found for a wide variety of functions.

Using the notation of integrals, we can express the probability of encountering a strategy x in the interval $[a, b]$. This is just the area under the curve $p(x)$ on this interval, or $\int_a^b p(x)\,dx$. But we saw that earlier that this is just the area from 0 to b minus the area from 0 to a (as in **Figure 2 (b)**) so,

$$\int_a^b p(x)\,dx = \int_0^b p(x)\,dx - \int_0^a p(x)\,dx = P(b) - P(a). \tag{3}$$

It is customary to use the symbol $P(x)\Big|_a^b$ to denote the difference $P(b) - P(a)$. So we would write

$$\int_a^b p(x)\,dx = P(x)\Big|_a^b.$$

In the case of situation 4, with $p(x) = 3x^2$ and $P(x) = x^3$, the probability that x lies in the interval $[a, b]$ is

$$\int_a^b 3x^2\,dx = x^3\Big|_a^b = b^3 - a^3.$$

sub For example, the probability that x lies in the interval $[0.2, 0.8]$ is

$$\int_a^b 3x^2\,dx = \int_{0.2}^{0.8} 3x^2\,dx = x^3\Big|_{0.2}^{0.8} = (0.8)^3 - (0.2)^3 = 0.512 - 0.008 = 0.504.$$

A more realistic situation:

Generally speaking, strategies are not restricted to finite intervals such as we have used in the previous examples. That is, x can take take on any nonnegative real value, $x \geq 0$. That means that a pdf $p(x)$ must be defined on the infinite interval $[0, +\infty)$, not just some finite interval such as $[0, 1]$ or $[0, 2]$. On the other hand, the area under this curve must still be 1 since it represents the probability of encountering some strategy, that is,

$$\int_0^\infty p(x)\,dx = 1.$$

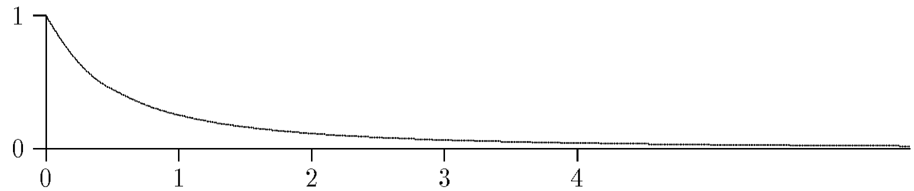
Such expressions are evaluated using limits. We first evaluate the expression $\int_0^b p(x)\,dx$. Then we see what happens to this expression as b gets infinitely large. The notation for this is

$$\lim_{b\to\infty} \int_0^b p(x)\,dx.$$

Here's a simple example: Suppose that the pdf was $p(x) = 1/(x+1)^2$ for $x > 0$. To show that this is a pdf, we need to show that the area under the curve (see **Figure 5**) is 1. That is we need to show that

$$\int_0^\infty \frac{1}{(x+1)^2}\,dx = 1.$$

Figure 5: The probability density function $p(x) = 1/(x+1)^2$ for $x \geq 0$. The total area under this infinitely long curve is 1.



It turns out that the cumulative distribution function (or in calculus terms, an antiderivative[1] for $p(x)$ is $P(x) = 1 - \frac{1}{x+1}$. Using this and limits to evaluate the integral above.

$$\int_0^\infty \frac{1}{(x+1)^2}\,dx = \lim_{b\to\infty}\int_0^b \frac{1}{(x+1)^2}\,dx = \lim_{b\to\infty} 1 - \frac{1}{x+1}\Big|_0^b \tag{4}$$

$$= \lim_{b\to\infty}\left(1 - \frac{1}{b+1}\right) - \left(1 - \frac{1}{0+1}\right). \tag{5}$$

Now as b gets large $\frac{1}{b+1}$ approaches 0. So

$$\int_0^\infty \frac{1}{(x+1)^2}\,dx = \lim_{b\to\infty}\left(1 - \frac{1}{b+1}\right) - \left(1 - \frac{1}{0+1}\right) = (1-0) - (1-1) = 1$$

as expected.

Our excursion to calculus is now complete and you may return return to the discussion of lifetime benefits at Prestwich's website.

[1] Calculus students note that the general antiderivative for $p(x) = \frac{1}{(x+1)^2} = (x+1)^{-2}$ is $P(x) = -(x+1)^{-1} + c$. We must choose c so that $P(0) = 0$, because the probability that an opponent quiting at cost less than 0 is 0.