Probability and Calculus (and Chicken)

One of the most useful parts of calculus in business is dealing with probability. Here's an example problem which we'll try to solve:

> Foster Industries runs a plant making stuffed chicken breast entrees for Sam's Club. The chicken entrees wholesale for \$5 per pound. The chicken breasts coming into the plant have an average size of 8 ounces. While the incoming chicken breasts vary in size around the average according to a probability distribution, the plant must produce an entree of fixed size. Chicken breasts too small to make an entree and any trimmings from chicken breasts which must be cut down to entree size can be sold as ground chicken at a wholesale price of \$1 per pound. What entree size maximizes the profitability of the plant?

To understand this question, we need to understand a bit about both probability and calculus. We can start thinking about this problem by imagining the two extreme cases: if all the chicken is ground, the average profit per breast is \$0.50. On the other hand, if the all the breasts weighed exactly 8 ounces, we could produce a single 8 oz entree from each breast with no waste, resulting in an average profit per breast of \$2.50.

1. Probability Density Functions, Mean, and Variance

Suppose we have a continuous random quantity X which can take on a range of values. The distribution of values of X is given by specifying a *probability density function* or pdf $p(x)$. We can answer natural questions about the variable X in terms of its pdf $p(x)$. For instance,

$$\text{The probability that } X \text{ is between } a \text{ and } b \text{ is } = \int_a^b p(x)\,dx.$$

Notice that the probability that X has a value *exactly equal* to a is the integral $\int_a^a p(x)\,dx$, which is 0 if $p(x)$ is a continuous function. Since the pdf is mostly integrated, it makes sense that an antiderivative for $p(x)$ would be a useful thing to have. The particular antiderivative we like is given by

$$c(x) = \int_{-\infty}^{x} p(t)\,dt.$$

Notice that $c(x)$ is the probability that X has a value less than x. This is called the *cumulative distribution function* or cdf of X. We can see a pdf and cdf in the pictures below.

¡figure¿

If a discrete random variable X has the value 10 with probability $1/2$ and the value 2 with probability $1/2$, we say that the *expected* value of X is given by $E(X) = 10 \times 1/2 + 2 \times 1/2 = 6$. This is the notion of "expected value" that we're all familiar with: if you have a 50% chance of winning \$20 playing cards, it should be (on average) "worth" \$10 to enter the game. We can extend this easily to continuous random variables:

$$\text{The expectation } E(X) \text{ or mean value } \mu \text{ of a random variable } X \text{ is } = \int_{-\infty}^{\infty} x\, p(x)\, dx.$$

It's a common mistake to think that a random variable is "likely" to have a value near the mean. This isn't always the case: the mean can be an average of larger values and smaller values, each of which can be even more common that the middle value, as in the picture below left. In the picture below right, we see two pdfs with the same mean, but very different shapes.

¡figure¿

We can measure how much a distribution is "bunched up" around the mean value μ by computing the *variance* of the distribution, which is given by

$$\text{The variance } V(X) \text{ of a random variable } X \text{ is } = \int_{-\infty}^{\infty} (x-\mu)^2\, p(x)\, dx.$$

that is, the variance is the mean of the squared difference between x and μ. The square root of variance is called the *standard deviation* of the random variable. Distributions with a small variance are tightly clustered around the mean value while distributions with large variance are quite spread out.

¡figure¿

2. The Normal Distribution

The single most important probability distribution is called the "normal distribution" or "Gaussian distribution". Its pdf $p(x)$ makes the famous "bell curve" shape. The pdf of a normal distribution with mean μ and variance σ^2 is given by

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

When you first see it, this distribution function looks completely weird. Why would *this* be the most important probability distribution function in the world? A complete answer to this question is somewhat beyond the scope of this class, but a partial answer is given by

Theorem 1 (Central Limit Theorem)**.** *Suppose that $X_1, \dots, X_n$ are a set of independent random variables with any probability distribution with mean μ and finite variance σ^2. Let S_n be the sample average*

$$S_n = \frac{1}{n}(X_1 + \cdots + X_n).$$

Then the distribution of the variable $\sqrt{n}(S_n - \mu)$ approaches a normal distribution with mean 0 and variance σ^2 ***regardless of the shape of the initial distribution****.*

This is a completely shocking result: if you average a sufficiently large collection of any kind of random variables, the results are *always* approximately normal! In fact, there's an even stronger version of the central limit theorem which applies to collections of random variables with different probability density functions as long the variances are all finite and the functions obey some technical conditions.

> **Example.** The height of American males in the 20-29 age range (according to the 1999 census) is quite close to a normal distribution with mean $\mu = 69.3$ inches and standard deviation $\sigma = 2.92$ inches (for females the mean is 64.1 and σ is 2.75). Why should this be so? Recent research (Nature Genetics 42, 565569 (2010)) shows that height is quite complicated, with the genetic proportion of height (mostly) explained by the combination of almost $295,000$ possible genetic differences (SNPs). If we assume that each of these differences exerts a positive or negative effect on height with a certain probability, the fact that the resulting heights are normally distributed is a consequence of the Central Limit Theorem!

The cumulative distribution function of the normal distribution is written in terms of the "error function" or $\mathrm{Erf}(x)$, which is defined by

$$\mathrm{Erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2}\, dt.$$

In fact, the cdf of the normal distribution with mean μ and standard deviation σ is given by

$$c(x) = \frac{1}{2}\left[1 + \mathrm{Erf}\left(\frac{x-\mu}{\sqrt{2\sigma^2}}\right)\right].$$

3. Foster Industries and the Chicken Problem

We now return to the chicken problem. We can reformulate the problem as follows:

> Suppose that the incoming chicken breasts are normally distributed with mean μ lbs and standard deviation σ lbs. Given a fixed entree size P, find the expected value of the profit per chicken breast as a function of P.

We can see for starters that there are two cases: the chicken breast size X is less than P (in which case the entire chicken breast is ground up) and the chicken breast size X is greater than P (in which case one entree is produced, plus some trimmings). Another observation that we can make here is that the *additional* profit for making an entree is always exactly $\$4P$, since we would have made $\$1P$ by selling that portion of the breast as ground chicken regardless. So the expected profit is given by

$$I(P) = 1 \times \int_{-\infty}^{\infty} p(x)\, dx + 4 \times P \int_P^{\infty} p(x)\, dx.$$