MthStat 465, Spring 2005, Lecture Number 17
Expected Value; The DeMoivre-Laplace Theorem and the Central Limit Theorem.

We define expected value, variance and standard deviation for a class of random variables. We then investigate how and when to approximate probability distributions by the standard normal distribution.

1. Expected value

For any random variable R with the properties

$$\sum_x \Pr(R = x) = 1$$
$$\sum_x |x| \Pr(R = x) < \infty$$

the expected value of R, denoted E[R], is defined by

$$\mathrm{E[R]} = \sum_{\mathrm{x}} \mathrm{x} \Pr(\mathrm{R} = \mathrm{x}) \tag{1}$$

and the variance of R, denoted by Var[R], is defined to be

$$\mathrm{Var[R]} = \mathrm{E[(R - E[R])^2]}. \tag{2}$$

It can be shown that for two random variables R and S and any constant c,

$$\mathrm{E[R + cS] = E[R] + cE[S]} \tag{3}$$

provided all the sums are convergent. This allows us to write

$$\mathrm{Var[R]} = \mathrm{E[R^2]} - (\mathrm{E[R]})^2 .$$

Furthermore, it is easy to show that

$$\mathrm{E[R^2]} = \sum_{\mathrm{x}} \mathrm{x}^2 \Pr(\mathrm{R} = \mathrm{x}). \tag{4}$$

We can now compute the expected value and variance of any binomial random variable easily. Consider the coin toss model. Let $I_j = 1$ if the j^{th} toss is a head, and $I_j = 0$ if it is a tail. In N tosses, the number of heads, R is equal to $I_1 + I_2 + \cdots I_N$. It is clear that if the probability of heads is p then $\mathrm{E[I_j] = p}$ and from (3) that $\mathrm{E[R] = Np}$. Since $R^2 = (I_1 + \cdots + I_N)^2$ we can find $\mathrm{E[R^2]}$ if we can find $\mathrm{E[I_j^2]}$ and $\mathrm{E[I_jI_k]}$ for $j \neq k$. In the former case, $I_j^2 = I_j$ so we get p again. In the latter case, the only time $I_jI_k \neq 0$ is when $I_j = I_k = 1$, that is, both trials give heads. This happens with probability p^2. Therefore $\mathrm{E[R^2] = Np + N(N-1)p^2}$, so

$$\mathrm{Var[R]} = \mathrm{E[R^2]} - (\mathrm{E[R]})^2 = \mathrm{Np(1-p)}.$$

The standard deviation of a random variable is the square root of its variance. In the case where

$$\sum_x \Pr(R = x) < 1$$

we can proceed as follows if the distribution function of R has a density, that is if there is some function f so that

$$\Pr(R \le x) = \int_{-\infty}^{x} f(u)\, du$$

for every x. In this case, if

$$\int_{-\infty}^{\infty} |u| f(u)\, du < \infty$$

then we define

$$\mathrm{E[R] := \int_{-\infty}^{\infty} uf(u)\, du.}$$

The general case of expected value is based on the case with sums. For more information, consult *Basic Probability Theory* by Robert Ash.

2. DeMoivre-Laplace Theorem

We saw last time that if

$$\Pr(R = x) = \binom{2N}{x} \left(\frac{1}{2}\right)^{2N}$$

for $x \in \{0, 1, \ldots, 2N\}$ then

$$\Pr(R = N + k) \approx \frac{1}{\sqrt{\pi}} \exp\left(\frac{k}{\sqrt{N}}\right)^2 \frac{1}{\sqrt{N}} \tag{5}$$

provided that k was an integer that was not too large. This suggests that

$$\sum_{k=-A}^{B} \Pr(R = N + k) \approx \int_{-A/\sqrt{N}}^{B/\sqrt{N}} \frac{1}{\sqrt{\pi}} \exp(-u^2)\, du. \tag{6}$$

We may write (6) a little more compactly as

$$\Pr(-A + N \leq R \leq B + N) \approx \int_{-A/\sqrt{N}}^{B/\sqrt{N}} \frac{1}{\sqrt{\pi}} \exp(-u^2)\, du. \tag{7}$$

In our case, the expected value of R is N and the variance of R is $N/2$. We wish to standardize our random variable by subtracting off its expected value and dividing the result by the standard deviation of the random variable. We get the following:

$$\begin{aligned} \Pr(-A + N \leq R \leq B + N) &= \Pr(-A \leq R - N \leq B) \\ &= \Pr\left(\frac{-A}{\sqrt{N/2}} \leq \frac{R - N}{\sqrt{N/2}} \leq \frac{B}{\sqrt{N/2}}\right) \\ &= \Pr\left(-a \leq \frac{R - N}{\sqrt{N/2}} \leq b\right) \end{aligned}$$

where we have put $a = A/\sqrt{N/2}$ and $b = B/\sqrt{N/2}$. Combining this with (7) gives us

$$\Pr\left(-a \leq \frac{R - N}{\sqrt{N/2}} \leq b\right) \approx \frac{1}{\sqrt{\pi}} \int_{-a/\sqrt{2}}^{b/\sqrt{2}} \exp(-u^2)\, du \tag{8}$$

and by letting $u = x/\sqrt{2}$ in the integral in (8),

$$\Pr\left(-a \leq \frac{R - N}{\sqrt{N/2}} \leq b\right) \approx \frac{1}{\sqrt{2\pi}} \int_{-a}^{b} \exp(-x^2/2)\, dx \tag{9}$$

In fact, what we have is equality in the limit:

$$\lim_{N\to\infty} \Pr\left(-a \le \frac{R-N}{\sqrt{N/2}} \le b\right) = \frac{1}{\sqrt{2\pi}} \int_{-a}^{b} \exp(-x^2/2)\, dx. \tag{10}$$

(10) is called the DeMoivre-Laplace Theorem.

3. The Central Limit Theorem

The DeMoivre-Laplace Theorem was is so handy that mathematicians wanted to extend it to other situations. The key ingredients are:

- The random quantity has a finite mean;
- The random quantity has a positive variance;
- The random quantity is the sum of smaller, non-interacting quantities.

This last property is a bit vague. In the coin tossing case, the number of heads is the sum of random variables I_k which are equal to 1 when a head is obtained on the k^{th} toss, and 0 otherwise. Since the tosses are independent, in some way these random variables do not influence one another. The key definition is the following. If $R_1, R_2, \ldots, R_N$ is a sequence of random variables, we say that they are an independent sequence if for any set of intervals $A_1, A_2, \ldots, A_N$,

$$\Pr(R_1 \in A_1, \ldots, R_N \in A_N) = \Pr(R_1 \in A_1) \times \cdots \Pr(R_N \in A_N).$$

In other words, random variables are independent if they always define independent events. The central limit theorem states:

Theorem 1 (Central Limit Theorem)**.** *If for for each positive integer N, $R_1, R_2, \ldots, R_N$ is a sequence of independent random variables with*

- $\Pr(R_k \le t) = \Pr(R_1 \le t)$ *for each real number t and each positive integer $k \le N$;*
- $\mathrm{E[R_k]} = \mathrm{m}$ *for each positive integer k;*
- $\sqrt{\mathrm{Var[R_k]}} = s > 0$ *for each positive integer k;*

then for each pair of real numbers $a < b$,

$$\lim_{N\to\infty} \Pr\left(a < \frac{R_1 + \cdots + R_N - Nm}{s\sqrt{N}} \le b\right) = \frac{1}{\sqrt{2\pi}} \int_{a}^{b} \exp(-u^2/2)\, du.$$

In the coin tossing model we take $R_k = I_k$. We have $\mathrm{E[I_k]} = \mathrm{p}$ and $\mathrm{Var[I_k]} = \mathrm{E[I_k^2]} - (\mathrm{E[I_k]})^2 = \mathrm{p} - \mathrm{p}^2 = \mathrm{p}(1-\mathrm{p})$. It can be shown that in the coin tossing model, where H denotes the number of heads in N tosses, the approximation

$$\Pr\left(a < \frac{H - Np}{\sqrt{Np(1-p)}} \le b\right) \approx \frac{1}{\sqrt{2\pi}} \int_{a}^{b} \exp(-u^2/2)\, du$$

is reasonably accurate so long as $p(1-p)N \ge 10$. One good trick for improving the accuracy is to compute $\Pr(A - 0.5 < H \le B + 0.5)$ instead of $\Pr(A \le H \le B)$ whenever A and B are integers.