Working with Probability Models

1. The Normal Model

We can use STATA to calculate similar values to those found in the Normal Table in the back of the book. Suppose we want to find the proportion of the area under the normal curve that lies below $z = 1$. To find this area we type

<p style="text-align:center; color:blue;">display normprob(1)</p>

in the command window. This gives us the result .84134475, which you can verify coincides with the value in the back of the book.

If, instead, we want to find the proportion of the area under the normal curve that lies above $z = 1$, we would need to type

<p style="text-align:center; color:blue;">display 1-normprob(1)</p>

Suppose we want to know how many standard deviations above the mean we need to be in order to lie in the 90th percentile of the normal curve. To find this value type

<p style="text-align:center; color:blue;">display invnorm(0.9)</p>

Using this command, we find that that the corresponding z-value is equal to 1.2815516. In other words, we need to be 1.28 standard deviations above the mean to be in the 90th percentile.

Ex. The height of U.S. men (in inches) approximately follows a normal model with mean 69.1 and standard deviation 2.9. Let X be the height of a randomly sampled man.

Suppose we want to estimate the probability that a man is between 5'6" and 6'. Then we can simply type

<p style="text-align:center; color:blue;">display normprob((72-69.1)/2.9)-normprob((66-69.1)/2.9)</p>

which gives us the result .69880214, or approximately 70% of all U.S. men are between 5'6" and 6'. Note that we, in this example, needed to insert the z-score into our calculations.

Suppose we want to know what is the shortest a man can be and still be in the top 10% of all U.S. males. We can calculate this value by first finding how many standard deviations, z, above the mean we need to be in order to be in the top 10%, and thereafter using the formula $x = z\sigma + \mu$ to find the proper value.

Doing these two tasks together we can write,

<p style="text-align:center; color:blue;">display invnorm(0.9)*2.9+69.1</p>

This gives us the result 72.8165 inches.

Exercise 1: The height of U.S. men (in inches) approximately follows a normal model with mean 69.1 and standard deviation 2.9. Let X be the height of a randomly sampled man.

    (a) Find the probability that a man is shorter than 60 inches.
    (b) Find the probability that a man is between 60 and 72 inches.
    (c) What is the shortest a man can be and still be in the top 20% of all U.S. males?

2. The Binomial Model

The STATA command Binomial(n,k,p) returns the probability of k or more successes in n trials when the probability of a success on a single trial is p.

If X is B(n,p), we can calculate $P(X \geq k)$ using STATA by typing

        display Binomial(n,k,p)

in the command window where n, k, and p are specified by the problem.

If we instead want to calculate $P(X = k)$ using STATA, we have to rewrite the equation as $P(X = k) = P(X \geq k) - P(X \geq k+1)$ (The probability of exactly k successes is equal to the probability of k or more successes minus the probability of k+1 or more successes). In STATA we can do this by typing

        display Binomial(n,k,p)-Binomial(n,k+1,p)

in the command window where n, k, and p are specified by the problem.

Ex. The probability that a baby is born a girl is .488. Suppose that 10 babies are born in one day at a certain hospital.

Let X = the number of girls. We can assume that X is bin(10, 0.488)

Find the probability that 5 or more of the babies are girls. Here n=10 and p=0.488. We want to calculate $P(X \geq 5)$, so k=5.

We use the command

        display Binomial(10,5,.488)

and find that the answer is .59318433.

Now suppose we want to calculate the probability that less than 5 of the infants are girls. Here n=10 and p=0.488. We want to calculate $P(X < 5) = 1 - P(X \geq 5)$.
We use the command

        display 1-Binomial(10,5,.488)

and find that the answer is .40681567.

If we want to calculate the probability that exactly 5 of the infants are girls, type

<p style="text-align:center; color:blue;">display Binomial(10,5,.488)-Binomial(10,6,.488)</p>

in the command window. Doing this we obtain the result .24538582.

Exercise 2: The probability that a baby is born a girl is .488. Suppose that 10 babies are born in one day in a hospital. What is the probability that

- (a) exactly three of the babies are girls?
- (b) three or more of the babies are girls?
- (c) less than three of the babies are girls?