MAT 211

Binomial and Normal Distributions

Binomial Distribution

The formula for a binomial probability is $B(n,p,k) = {}_nC_k(p)^k(1-p)^{n-k}$, where p = probability of success, n = number of trials, and k = number of successes. Using the TI-83 calculator, $B(n,p,k) = binompdf(n,p,k)$. Essentially, this finds the probability of exactly k successes out of n Bernoulli Trials, when the probability of success is p.

We can also find the probability of *at most* k successes out of n Bernoulli Trials by using $binomcdf(n,p,k)$.

Example: Stan is a target shooter. He has a probability of 0.85 of making his shot (assume independence). He makes 30 attempts at the target. Define X to be the number of successful shots. Find the following probabilities:

a) $P(X = 20) = binompdf(30,.85,20) = .0067152$
b) $P(X \le 22) = binomcdf(30,.85,22) = .069784$
c) $P(X \ge 24) = 1 - binomcdf(30,.85,23) = binomcdf(30,.15,6) = .84741$
(**Note:** Getting at least 24 successes is the compliment of at most 23 successes. Also, getting at least 24 successes is the exact same thing as getting at most 6 failures.)
d) $P(20 \le X \le 26) = binomcdf(30,.85,26) - binomcdf(30,.85,19) = .6754$

Normal Distribution

In statistics, the normal distribution curve (also known as the bell curve) is a model that is typically used in situations where the data tends to collect near the mean average and thin out further away from the mean. The normal curve is designed to have the mean at $\mu = 0$ and for about 68% of the data to be within 1 standard deviation (σ) of the mean (so, 34% of the data is one standard deviation above and 34% is one standard deviation below), for 95% of the data to be within 2 standard deviations of the mean, and for over 99% of the data to be within 3 standard deviations of the mean.

Since normal distribution is a continuous situation (rather than discrete), it is virtually impossible to find the probability of an exact data point. So, instead we calculate probabilities over an interval of the data. Since we view these probabilities as 'areas under the curve' we can typically use calculus (antiderivatives) to find these values. However, the function used to create the graph of the normal curve is quite difficult to integrate, so we can use our calculator instead. Once again, the TI-83 has the capabilities to find these probabilities. The probability that a data point falls in the interval $[a,b]$ can be found by using $normalcdf(\sigma_1,\sigma_2)$, where σ_1 is the number of standard deviations that the value 'a' is away from the mean and σ_2 is the number of standard deviations that the value 'b' is away from the mean.

<u>Example</u>: We are told that the mean IQ of a population is $\mu = 100$ and the standard deviation is $\sigma = 10$. So, 68% of the population has an IQ between 90 and 110, 95% of the population has an IQ between 80 and 120, and over 99% of the population has an IQ between 70 and 130. We can verify these values by using *normalcdf* on our calculators. So, $P(90 \le X \le 110) = normalcdf(-1,1) = .68268$ since 90 is -1 standard deviation from the mean (ie. it is 1 standard deviation *less than* the mean) and 110 is +1 standard deviation from the mean (ie. it is 1 standard deviation *greater than* the mean). Similarly, $P(80 \le X \le 120) = normalcdf(-2,2) = .95449$ and $P(70 \le X \le 130) = normalcdf(-3,3) = .9973$. Sometimes, it is not very obvious how many standard deviations from the mean (what is known as the z-score) a data point actually is. In these cases we use a technique called normalization to find the z-score. To find the z-score we use the following formula: $z = \frac{x - \mu}{\sigma}$, where x is the value of the data point, μ is the mean, and σ is the standard deviation. Typically, the smallest z-score that we need to use is -5 and the greatest z-score is +5.

a) What is the probability that a randomly selected person has an IQ between 84 and 102?

First we find the corresponding z-scores. The z-score for 84 is $\frac{84-100}{10} = -1.6$ and the z-score for 102 is $\frac{102-100}{10} = .2$. So, $P(84 \le X \le 102) = normalcdf(-1.6,.2) = .52446$.

b) What is the probability that a randomly selected person has an IQ less than 115?

The z-score for 115 is $\frac{115-100}{10} = 1.5$ and the z-score for the low end is -5.

So, $P(X \le 115) = normalcdf(-5,1.5) = .93319$.

c) What IQ interval covers the lower 36% of the population?

To solve this we can use a function on the calculator called *InvNorm*. *InvNorm*(.36) outputs the z-score, z, for the value such that 36% of the area under the normal curve is in the interval $X \le z$. In this case, $z = InvNorm(.36) = -.3585$. To find the IQ score associated with this z-value, we can use the normalization formula:

$-.3585 = \frac{x-100}{10}$ => $x = -.3585(10) + 100 = 96.415$. So, 36% of the population has an IQ less than or equal to 96.415.

d) The High IQ Society only accepts people with an IQ in the top 2%. What is the minimum IQ needed to be accepted in the society?

InvNorm only finds z-value associated with area to the left of that value. So, we must find $z = InvNorm(.98) = 2.053$. So, $2.053 = \frac{x-100}{10}$ => $x = 2.053(10) + 100 = 120.53$. So, a person must have an IQ higher than 120.53 to be in the Society.