Chapter Five

Normal Distribution: Suppose μ and σ are given parameters. Consider the function

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

for x any real number.

Then, f(x) is a pdf and the resulting distribution is called the <u>normal distribution</u> denoted $N(\mu, \sigma^2)$. This is also called the "Bell Curve". Notice how the pdf is symmetric with respect to the y-axis.

Result: The mean and standard deviation of the Normal Distribution are precisely the parameters supplied at the beginning of the problem.

Remark: To obtain a closed form for the integral of f(x) is impossible. However, to obtain probabilities by formulas, this is precisely what we must do. Making a table of integral values using numerical methods (Trapezoidal Rule, Simpson's Rule, Gaussian Quadrature) proves very useful. However, this table will be different for each selection of parameters μ and σ.

Standard Normal Distribution: Consider the normal distribution N(0,1). Table values for the resulting distribution function F(z), called the <u>standard normal distribution</u>, are given in the Appendix in Table IV as well as on many statistical calculators where one computes F(x) = normalcdf(μ, σ, x). Notice, when using the standard normal distribution, we will always use the random variable z instead of using x.

Result: Probabilities in $N(\mu, \sigma^2)$ can always be converted to probabilities in N(0,1) by using the normalizing change of variables

$$Z = (x - \mu)/\sigma.$$

Result: In N(0,1) if c<0, then F(c) = 1 - F(-c). So, whenever using F(b) - F(a) and one of a or b is negative, apply this result to switch the negative value to a positive table-supplied value.

Theorem: If X is $N(\mu,\sigma^2)$, then the random variable $V = Z^2 = (X-\mu)^2/\sigma^2$ is $\chi^2(1)$.

Pf: Since $v \geq 0$, $G(v) = P(V \leq v) = P(Z^2 < \sqrt{v}\ ) = P(-\sqrt{v} < Z < \sqrt{v}\ )$

$$= 2\int_0^{\sqrt{v}} \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz$$

$$= 2\int_0^{v} \frac{1}{\sqrt{2y\pi}} e^{-y/2} dy$$

using the change of variable $z^2 = y$. By the FTOC, we get for $0 < v < \infty$

$$g(v) = G'(v) = \frac{1}{\sqrt{2v\pi}} e^{-v/2} = \frac{v^{1/2-1}}{\sqrt{2\pi}} e^{-v/2}.$$

Since G(v) is a distribution function, then g(v) must be a pdf and so

$$\int_0^{\infty} \frac{v^{1/2-1}}{\sqrt{2\pi}} e^{-v/2} dv = 1$$

Letting x = v/2 yields

$$\frac{1}{\sqrt{\pi}} \Gamma(\tfrac{1}{2}) = \frac{1}{\sqrt{\pi}} \int_0^{\infty} x^{1/2-1} e^{-v/2} dv = 1$$

Hence, $\Gamma(1/2) = \sqrt{\pi}$ and so g(v) is the pdf for $\chi^2(1)$.

HOMEWORK: page 277

Theorem: Let $X_1, X_2, \ldots, X_n$ be a random sample from a normal distribution with mean μ and variance σ^2. Then, the random variable $\overline{X} = \sum_{k=1}^{n} X_k / n$ is normally distributed with mean μ and variance σ^2/n.

Theorem: (Central Limit Theorem):

If Y is the mean of a random sample $X_1, X_2, \ldots, X_n$ from a distribution with a finite mean μ and a finite positive variance σ^2, then, the distribution of W = $\frac{\overline{x}-\mu}{\sigma/\sqrt{n}}$ becomes a normal variable in the limit as n approaches infinity.

Alternatively, $\overline{x} \approx N(\mu, \sigma^2)$.

Normal Approximation to Binomial: Let $X_1, X_2, \ldots, X_n$ be Bernoulli variables; $Y = X_1 + X_2 + \ldots + X_n$ then is Binomial. For each Bernoulli variable, $\mu = p$ and $\sigma^2 = p(1-p)$.

By the CLT, $\overline{x}$ = Y/n is approximately normally distributed. So,

$$W = (\overline{x} - p) / \sqrt{np(1-p)} = (Y - np) / \sqrt{np(1-p)} = Z,$$

the standard normal variable. That is, the binomial distribution can be approximated by N(p, np(1-p)).

Be sure to expand the intervals to convert (discrete) pdf values to (continuous) areas of histograms.

This approximation is generally ok provided $np \geq 5$ and $n(1-p) \geq 5$.

Normal Approximation to Poisson: Since for large n, the binomial and Poisson distributions are very close, replace the Binomial mean np with the Poisson mean $\mu = \lambda T$ and the Binomial variance np(1-p) with the Poisson variance $\sigma^2 = \lambda T = \mu$. Therefore

$$W = \frac{Y - \lambda T}{\sqrt{\lambda T}}$$

is also approximately standard normal.

HOMEWORK: page 303