Lecture 11 ctd

Confidence intervals for variance.

The point estimate of variance is:

$$\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

The distribution of :

$\frac{\hat{\sigma}^2}{\sigma^2}(n-1)$ has been worked out. It is called the Chi-Square distribution on n-1 degrees of freedom.
Inconsistently CHIDIST(x,df) = P(X>x)

In fact it is a special case of the Gamma distribution.

$$\mathrm{X}_{n-1}^2 = Gamma\left(\frac{n}{2},\frac{1}{2}\right).$$

Like the t it has one parameter and in tables the percentage points are tabulated.

To construct, say, a 95% confidence interval we look up the values such that:

$$P(X_{n-1}^2 > U) = \frac{\alpha}{2} \quad \text{and} \quad P(X_{n-1}^2 < L) = \frac{\alpha}{2}$$

$$U = CHIINV(\frac{\alpha}{2}, n-1) \quad L = CHIINV(1-\frac{\alpha}{2})$$

The function CHIINV(α, df) returns the value so that there is a probability of α to the ***right*** of this.

$$P(L < \frac{\hat{\sigma}^2}{\sigma^2}(n-1) < U) = 1-\alpha$$

Rearranging gives

$$\frac{1}{L} > \frac{\sigma^2}{\hat{\sigma}^2(n-1)} > \frac{1}{U}$$

$$\frac{(n-1)\hat{\sigma}^2}{L} > \sigma^2 > \frac{(n-1)\hat{\sigma}^2}{U}$$

gives the confidence interval.

If you want standard deviation just take the square root of the limits.

Example:
20 observations give an estimate of variance $\hat{\sigma}^2 = 5.6$.

$$\hat{\sigma} = 2.37$$

The 95% confidence interval for the variance is:

L=CHIINV(0.975,19)=8.907

U=CHIINV(0.025,19)=32.852

$$\frac{5.6*19}{8.907} > \sigma^2 > \frac{5.6*19}{32.852}$$

$$11.95 > \sigma^2 > 3.24$$

$3.46 > \sigma > 1.80$ for standard deviation.

So quite a big range the standard deviation could be 50% larger or 50% smaller than the estimated value.

The chi-square can be used for testing whether the variance is a given value. (Not a common thing to want to do).

It also arises in a different type of test but this will (probably) not be covered in this course.