Math 3215: Lecture 18

Will Perkins

March 15, 2012

1 More on Estimators

Let $X_1, \dots X_n$ be iid, $N(\mu, \sigma^2)$.

- Find unbiased estimators for both μ and σ^2.
- Are these also maximum likelihood estimators?
- What is the variance of the estimators?

2 Confidence Intervals

Last class we saw how to estimate the probability that the true value of a parameter p was within a given error of our estimate. A *confidence interval* gives a range of values so that with a certain probability (or confidence) the true value of the paramter is in that range. Specifically, if:

$$\Pr[\hat{\Theta} - t \le p \le \hat{\Theta} + s] = .95$$

Then $[\hat{\Theta} - t, \hat{\Theta} + s]$ is a 95% confidence interval for p. We almost always use normal distributions (or other closely related distributions) to calculate confidence intervals.

2.1 Confidence Intervals with known variance

If we assume (unrealistically) that we know the variance of our random variable, but do not know the mean, the the construction of a confidence interval is straightforward. Steps:

1. Start with the corresponding confidence interval for Z, a standard normal. Eg. $\Pr[-2 \le Z \le 2] = .95$.
2. With a series of steps of algebra, transform the above confidence interval into one for your estimator $\hat{\Theta}$, under the assumption from the CLT that $\hat{\Theta}$ is approximately normal.

Examples:

Some useful facts about the Normal Distribution

- $\Pr[-1 \le Z \le 1] \sim .68$
- $\Pr[-2 \le Z \le 2] \sim ..95$
- $\Pr[-3 \le Z \le 3] \sim .99$

2.2 Confidence intervals with unknown variance

A more realistic situation is that we don't know the mean or variance of our underlying distribution.

- What would you do in that case?

2.2.1 Student's t-distribution

Let $X_1, \dots X_n$ be iid with mean μ and variance σ^2. The sample mean, $\overline{X}$, is our MLE for the unknown μ. The CLT tells us:

$$\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \sim Z$$

where $Z \sim N(0, 1)$.

But we are interested in the case that we don't know σ. William Gosset, working for Guiness Brewery and publishing under the pseudonym 'Student' proved the following theorem:

$$\frac{\overline{X} - \mu}{s/\sqrt{n}} \sim T_n$$

where s^2 is the unbiased estimator of σ^2 we calculated above. (i.e. $s^2 = \frac{\sum (X_i - \overline{X})^2}{n}$) and T_n is a distribution called Student's T disribution with $n - 1$ degrees of freedom. Don't worry too much about the degrees of freedom. You can look up values of T on the computer on in a table, but the main point is that as n gets large (say > 20), T_n is very close to Z.