Confidence Intervals (3 pages; 16/4/16)

(1) Mean of a Normal distribution with known variance

$X \sim N(\mu, \sigma^2) \Rightarrow \overline{X} \sim N(\mu, \frac{\sigma^2}{n})$

$\Rightarrow \mu - 1.96\frac{\sigma}{\sqrt{n}} < \overline{x} < \mu + 1.96\frac{\sigma}{\sqrt{n}}$ with 95% probability

Then, for given $\overline{x}$, $\mu < \overline{x} + 1.96\frac{\sigma}{\sqrt{n}}$ & $\mu > \overline{x} - 1.96\frac{\sigma}{\sqrt{n}}$;

ie $\overline{x} - 1.96\frac{\sigma}{\sqrt{n}} < \mu < \overline{x} + 1.96\frac{\sigma}{\sqrt{n}}$ in 95% of cases (note that

μ is fixed: we shouldn't say that it lies within the above interval with 95% probability - the probability is either 0 or 1 for a particular $\overline{x}$)

Thus a 95% confidence interval for μ is

$(\overline{x} - 1.96\frac{\sigma}{\sqrt{n}}, \overline{x} + 1.96\frac{\sigma}{\sqrt{n}})$

(assuming that a symmetrical interval is required).

(2) Mean of a Normal distribution with unknown variance, where the sample is large

As the sample is large (usually taken to be ≥ 30), s (based on a divisor of $n-1$) can be assumed to be a reasonably good approximation to σ, so that a 95% confidence interval for μ is $(\overline{x} - 1.96\frac{s}{\sqrt{n}}, \overline{x} + 1.96\frac{s}{\sqrt{n}})$

(3) Mean of an unknown distribution (with unknown variance), where the sample is large

As the sample size is large, the Central Limit theorem says that $\overline{X}$ approx. $\sim N(\mu, \frac{\sigma^2}{n})$ [$n \geq 30$ also applies here] and s can be assumed to be a reasonably good approximation to σ (again, as the sample size is large).

A 95% confidence interval for μ is then $(\bar{x} - 1.96\frac{s}{\sqrt{n}}, \bar{x} + 1.96\frac{s}{\sqrt{n}})$; ie as for (2) (so, for practical purposes, it doesn't matter whether the distribution is Normal or not).

(4) Mean of a Normal distribution with unknown variance, where the sample is small

To reflect the greater uncertainly caused by approximating σ by s when the sample is small, the t-distribution is used, with

$\upsilon = n - 1$ degrees of freedom.

Eg, for $\upsilon = 20$, a 95% confidence interval for μ is

$(\bar{x} - 2.086\frac{s}{\sqrt{n}}, \bar{x} + 2.086\frac{s}{\sqrt{n}})$;

(Note that the underlying distribution has to be Normal, in order for the t-distribution to apply. As the sample size increases, the t-value tends to the z-value.)

(5) Binomial proportion, for a large sample (using a Normal approximation)

Let $X \sim B(n, p)$. If n is large and p is not too small, in such a way that a Normal approximation is appropriate (this will usually be the case if $n \geq 50$ & $np \geq 10$), then X approx. $\sim N(np, np(1-p))$.

If $Y = \frac{X}{n}$ is the proportion of successes, then

$E(Y) = \frac{1}{n}E(X) = \frac{np}{n} = p$

and $Var(Y) = \frac{1}{n^2}Var(X) = \frac{np(1-p)}{n^2} = \frac{p(1-p)}{n}$

and so Y approx. $\sim N(p, \frac{p(1-p)}{n})$.

If $\hat{p}$ is the observed proportion of successes in n trials, and π is the population proportion, then a 95% confidence interval for π is

$(\hat{p} \ - 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \hat{p} + 1.96\sqrt{\frac{\hat{p}(1-\hat{p})}{n}})$

[by the same reasoning as in (1)]

(6) Mean of a Poisson distribution (using a Normal approximation)

$X \sim Po(\lambda)$ approx. $\sim N(\lambda, \lambda)$

If $\hat{\lambda}$ is the observed value of X [$\hat{\lambda}$ should be sufficiently large to justify the Normal approximation; generally ≥ 20], then a 95% confidence interval for λ is $(\hat{\lambda} - 1.96\sqrt{\hat{\lambda}}, \hat{\lambda} + 1.96\sqrt{\hat{\lambda}})$

(7) Variance/standard deviation of a Normal distribution

If $X \sim N(\mu, \sigma^2)$ and s^2 is obtained from a sample of size n, then it can be shown that

$\chi^2_{n-1}(0.025) < \frac{(n-1)s^2}{\sigma^2} < \chi^2_{n-1}(0.975)$ with 95% probability,

so that a 95% confidence interval for σ^2 is $(\frac{(n-1)s^2}{\chi^2_{n-1}(0.975)}, \frac{(n-1)s^2}{\chi^2_{n-1}(0.025)})$

and a 95% confidence interval for σ is $(\frac{s\sqrt{n-1}}{\sqrt{\chi^2_{n-1}(0.975)}}, \frac{s\sqrt{n-1}}{\sqrt{\chi^2_{n-1}(0.025)}})$