Inferences for population variances

Sampling distribution of S^2

Suppose $Y \sim \text{normal}(\mu, \sigma^2)$, Y_1, Y_2, ..., Y_n is a random sample, and

$$S^2 = \frac{1}{(n-1)}\left[(Y_1 - \overline{Y})^2 + (Y_2 - \overline{Y})^2 + \cdots + (Y_n - \overline{Y})^2\right]$$

is the sample variance. Then

$$\frac{(n-1)}{\sigma^2} S^2 \sim \text{chi-square}(n-1)$$

(chi-square distribution with $n-1$ degrees of freedom).
Properties:

$$\mathrm{E}\left(\frac{(n-1)}{\sigma^2} S^2\right) = n - 1$$

and so

$$\mathrm{E}\left(S^2\right) = \sigma^2 \quad (S^2 \text{ is an unbiased estimate})$$

Data: $n,\ \overline{y},\ s^2$

$100(1-\alpha)\%$ CI for σ^2

$$\left(\frac{(n-1)s^2}{\chi^2_{\alpha/2}},\ \frac{(n-1)s^2}{\chi^2_{1-(\alpha/2)}}\right) = (l,\ u)$$

Also, $\left(\sqrt{\cdot},\ \sqrt{\cdot}\right)$ is a $100(1-\alpha)\%$ CI for σ

Hypothesis tests

H_0: $\sigma^2 = \sigma_0^2$ (known constant)

$$\text{H}_\text{a}:\ \sigma^2 \left\{ \begin{array}{c} > \\ < \\ \neq \end{array} \right\} \sigma_0^2$$

Test statistic:

$$\chi^2 = \frac{(n-1)s^2}{\sigma_0^2}$$

Rejection region:

$$\text{reject H}_0 \text{ if } \chi^2 \left\{ \begin{array}{c} \geq \chi^2_\alpha \\ \leq \chi^2_{1-\alpha} \\ \left(\geq \chi^2_{\alpha/2}\right) \text{ or } \left(\leq \chi^2_{1-(\alpha/2)}\right) \end{array} \right\}$$

The F distribution

If $U \sim$ chi-square(j) and $V \sim$ chi-square(k), and U and V are independent, then

$$F = \frac{(U/j)}{(V/k)}$$

has an F distribution (named after R. A. Fisher) with j, k degrees of freedom. We write $F \sim \mathrm{F}(j, k)$.

Two populations: one is normal$(\mu_1, {\sigma_1}^2)$, the other is normal$(\mu_2, {\sigma_2}^2)$. From each a random sample will be taken, yielding $n_1, \overline{Y}_1, {S_1}^2$; $n_2, \overline{Y}_2, {S_2}^2$. Then

$$F = \frac{\frac{1}{(n_1-1)} \frac{(n_1-1)}{{\sigma_1}^2} {S_1}^2}{\frac{1}{(n_2-1)} \frac{(n_2-1)}{{\sigma_2}^2} {S_2}^2} = \frac{({S_1}^2/{S_2}^2)}{({\sigma_1}^2/{\sigma_2}^2)} \sim \mathrm{F}(n_1 - 1, n_2 - 1)$$

Note: property of the F distribution is that

$$f_{1-p}^{(j,k)} = \frac{1}{f_p^{(k,j)}}$$

Data: $n_1, \overline{y}_1, {s_1}^2;\ n_2, \overline{y}_2, {s_2}^2$

100(1 − α)% CI for $\frac{{\sigma_1}^2}{{\sigma_2}^2}$

$$\left(\frac{{s_1}^2}{{s_2}^2} f_{1-(\alpha/2)}, \frac{{s_1}^2}{{s_2}^2} f_{\alpha/2}\right) = (u, l)$$

Hypothesis test (Bartlett's test)

$$\text{H}_0\text{: } \frac{{\sigma_1}^2}{{\sigma_2}^2} = 1 \qquad ({\sigma_1}^2 = {\sigma_2}^2)$$

$$\text{H}_\text{a}\text{: } \frac{{\sigma_1}^2}{{\sigma_2}^2} \left\{ \begin{array}{c} > \\ \neq \end{array} \right\} 1 \qquad \left(\begin{array}{c} {\sigma_1}^2 > {\sigma_2}^2 \\ {\sigma_1}^2 \neq {\sigma_2}^2 \end{array} \right)$$

Test statistic:

$$f = \frac{{s_1}^2}{{s_2}^2}$$

Rejection region:

$$\text{reject H}_0 \text{ if } f \left\{ \begin{array}{c} \geq f_\alpha \\ \left(\leq f_{1-(\alpha/2)} \right) \text{ or } \left(\geq f_{\alpha/2} \right) \end{array} \right\}$$

Note:

χ^2 and F-tests for variances are **not** robust to departures of populations from normality.

t-tests for means **are** robust to (moderate) normality departures.