Introduction to Probability and Statistics

Review of interval estimates

April 3, 2009

Here is a list of all of the situations for which we have constructed confidence intervals. In each case the situation is that we have drawn a random sample $X_1, \ldots, X_n$ from a population whose distribution is known up to some parameter.

Normal population, mean μ unknown, variance σ^2 known

- **Unbiased estimator for** μ: $\bar{X} = \frac{X_1 + \ldots + X_n}{n}$
- **Maximum likelihood estimator for** μ: $\bar{X}$
- $100(1-\alpha)\%$ **confidence intervals for** μ:
  - **Two-sided**: $\left(\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right)$
  - **One-sided lower**: $\left(-\infty, \bar{X} + z_{\alpha}\frac{\sigma}{\sqrt{n}}\right)$
  - **One-sided upper**: $\left(\bar{X} - z_{\alpha}\frac{\sigma}{\sqrt{n}}, \infty\right)$
- **All based on**: $\frac{\bar{X}-\mu}{\sigma/\sqrt{n}} = Z$, a standard normal

Normal population, mean μ unknown, variance σ^2 unknown, estimating mean

- $100(1-\alpha)\%$ **confidence intervals for** μ:
  - **Two-sided**: $\left(\bar{X} - t_{\alpha/2,n-1}\frac{S}{\sqrt{n}}, \bar{X} + t_{\alpha/2,n-1}\frac{S}{\sqrt{n}}\right)$
  - **One-sided lower**: $\left(-\infty, \bar{X} + t_{\alpha,n-1}\frac{S}{\sqrt{n}}\right)$
  - **One-sided upper**: $\left(\bar{X} - t_{\alpha}\frac{S}{\sqrt{n}}, \infty\right)$
- **All based on**: $\frac{\bar{X}-\mu}{S/\sqrt{n}} = t_{n-1}$, a t-distribution with $n-1$ degrees of freedom, where $S^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}$ is sample variance

Normal population, mean μ unknown, variance σ^2 unknown, estimating variance

- **Unbiased estimator for** σ^2: $S^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}$
- **Maximum likelihood estimator for** σ^2: $\frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n}$
- $100(1-\alpha)\%$ **confidence intervals for** σ^2:
  - **Two-sided**: $\left(\frac{(n-1)S^2}{\chi^2_{\alpha/2,n-1}}, \frac{(n-1)S^2}{\chi^2_{1-\alpha/2,n-1}}\right)$
  - **One-sided lower**: $\left(0, \frac{(n-1)S^2}{\chi^2_{1-\alpha,n-1}}\right)$
  - **One-sided upper**: $\left(\frac{(n-1)S^2}{\chi^2_{\alpha,n-1}}, \infty\right)$
- **All based on**: $\frac{(n-1)S^2}{\sigma^2} = \chi^2_{n-1}$, a χ^2 distribution with $n-1$ degrees of freedom

Two normal populations, means μ_1, μ_2 unknown, variances σ_1^2, σ_2^2 known, estimating difference between means

- $100(1-\alpha)\%$ **confidence intervals for** $\mu_1 - \mu_2$:
  - **Two-sided**: $\left(\bar{X} - \bar{Y} - z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}, \bar{X} - \bar{Y} + z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}\right)$, where $\bar{X}$ is sample mean from first population (sample size n), $\bar{Y}$ is sample mean from second population (sample size m)
  - **One-sided lower**: $\left(-\infty, \bar{X} - \bar{Y} + z_{\alpha}\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}\right)$
  - **One-sided upper**: $\left(\bar{X} - \bar{Y} - z_{\alpha}\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}, \infty\right)$
- **All based on**: $\frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n} + \frac{\sigma_2^2}{m}}} = Z$, a standard normal

Two normal populations, means μ_1, μ_2 unknown, variances σ_1^2, σ_2^2 unknown, assumed equal to σ^2, estimating difference between means

- **Pooled estimator for common variance**: $S_p^2 = \frac{(n-1)S_1^2+(m-1)S_2^2}{n+m-2}$
- $100(1-\alpha)\%$ **confidence intervals for** $\mu_1 - \mu_2$:
  - **Two-sided**: $\left(\bar{X} - \bar{Y} - t_{\alpha/2,n+m-2}\sqrt{\frac{S_p^2}{n} + \frac{S_p^2}{m}}, \bar{X} - \bar{Y} + t_{\alpha/2,n+m-2}\sqrt{\frac{S_p^2}{n} + \frac{S_p^2}{m}}\right)$
  - **One-sided lower**: $\left(-\infty, \bar{X} - \bar{Y} + t_{\alpha,n+m-2}\sqrt{\frac{S_p^2}{n} + \frac{S_p^2}{m}}\right)$
  - **One-sided upper**: $\left(\bar{X} - \bar{Y} - t_{\alpha,n+m-2}\sqrt{\frac{S_p^2}{n} + \frac{S_p^2}{m}}, \infty\right)$
- **All based on**: $\frac{(\bar{X}-\bar{Y})-(\mu_1-\mu_2)}{\sqrt{\frac{S_p^2}{n}+\frac{S_p^2}{m}}} = t_{n+m-2}$, a t-distribution with $n+m-2$ degrees of freedom

Bernoulli population, parameter p unknown

Here $X_1, \ldots, X_n$ is a sample of size n from a population in which an unknown proportion p have a certain characteristic, and each X_i is 1 if sample i has the characteristic, and 0 otherwise.

- Approximate $100(1-\alpha)\%$ **confidence intervals for** p**, if sample size is large (**$n \geq 30$**)**:
  - Using $\hat{p} = \frac{X_1+\ldots+X_n}{n}$ as an estimate for p in the variance:
    * **Two-sided**: $\left(\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right)$
    * **One-sided lower**: $\left(0, \hat{p} - z_{\alpha}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right)$
    * **One-sided upper**: $\left(\hat{p} - z_{\alpha}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, 1\right)$
  - Using worst-case estimate $p = 1/2$ in the variance:
    * **Two-sided**: $\left(\hat{p} - z_{\alpha/2}\sqrt{\frac{1}{4n}}, \hat{p} + z_{\alpha/2}\sqrt{\frac{1}{4n}}\right)$
    * **One-sided lower**: $\left(0, \hat{p} - z_{\alpha}\sqrt{\frac{1}{4n}}\right)$
    * **One-sided upper**: $\left(\hat{p} - z_{\alpha}\sqrt{\frac{1}{4n}}, 1\right)$
- **All based on**: $\frac{\bar{X}-np}{\sqrt{np(1-p)}} \approx Z$, a standard normal (Central Limit Theorem), where $X = X_1 + \ldots + X_n$.