Classical and Bayesian inference

AMS 132

February 21, 2018

Confidence Intervals for the Mean of a Normal Distribution

- Confidence intervals provide a method of adding more information to an estimator $\widehat{\theta}$ when we wish to estimate an unknown parameter θ. We can find an interval (A, B) that we think has high probability of containing θ. The length of such an interval gives us an idea of how closely we can estimate θ.

Example

Suppose that $X_1, \ldots, X_n$ form a random sample from the normal distribution with mean μ and variance σ^2. We know that an estimator for μ is $\overline{X}_n$, but how close is it from μ? How much confidence should we place in the estimator $\overline{X}_n$ for μ?

Confidence Intervals for the Mean of a Normal Distribution

a) Suppose that $X_1, \ldots, X_n$ form a random sample from the normal distribution with mean μ and variance σ^2.

b) Consider $\overline{X}_n$ and σ' estimators for μ and σ, respectively. What is the distribution of $U = \frac{\sqrt{n}(\overline{X}_n - \mu)}{\sigma'}$?

c) Find c such that $P(-c < U < c) = 0.95$. More generally, Find c such that $P(-c < U < c) = \gamma$, $0 < \gamma < 1$.

d) Use c) to find the random variables A and B, such that $P(A < \mu < B) = 0.95$.

Confidence Intervals for the Mean of a Normal Distribution

Definition (Confidence Interval)

Let $\boldsymbol{X} = (X_1, \ldots, X_n)$ be a random sample from a distribution that depends on a parameter (or parameter vector) θ. Let $g(\theta)$ be a real-valued function of θ. Let $A \leq B$ be two statistics that have the property that for all values of θ,

$$Pr(A < g(\theta) < B) \geq \gamma.$$

Then the random interval (A, B) is called a *coefficient* γ *confidence interval for* $g(\theta)$ or a 100γ *percent confidence interval for* $g(\theta)$.

If the inequality "$\geq \gamma$" is replaced by an equality for all θ, the confidence interval is called *exact*.

After the values of the random variables $X_1, \ldots, X_n$ in the random sample have been observed, the values of $A = a$ and $B = b$ are computed, and the interval (a, b) is called the *observed value of the confidence interval*.

Confidence Intervals for the Mean of a Normal Distribution

Theorem (Confidence Interval for the Mean of a Normal Distribution)

Let $X_1, \ldots, X_n$ be a random sample from the normal distribution with mean μ and variance σ^2. For each $0 < \gamma < 1$, the interval (A, B) with the following endpoints is an exact coefficient γ confidence interval for μ:

$$A = \overline{X}_n - T_{n-1}^{-1}\left(\frac{1+\gamma}{2}\right)\frac{\sigma'}{\sqrt{n}},$$

$$B = \overline{X}_n + T_{n-1}^{-1}\left(\frac{1+\gamma}{2}\right)\frac{\sigma'}{\sqrt{n}}.$$

Confidence Intervals for the Mean of a Normal Distribution

Example (Lactic acid concentration in cheese)

One chemical whose concentration can affect taste is lactic acid. Cheese manufacturers who want to establish a loyal customer base would like the taste to be about the same each time a customer purchases the cheese. The variation in concentrations of chemicals like lactic acid can lead to variation in the taste of cheese.

Suppose that we model the concentration of lactic acid in several chunks of cheese as independent normal random variables with mean μ and variance σ^2.

a) Find a 90 percent confidence interval for μ, the unknown mean lactic acid concentration.
b) Suppose that we observe the following lactic acid concentrations $0.86, 1.53, 1.57, 1.81, 0.99, 1.09, 1.29, 1.78, 1.29, 1.58$. Compute the observed value of the 90 percent confidence interval for μ.
c) Make an interpretation of the intervals found in a) and b).

Confidence Intervals for the Mean of a Normal Distribution: interpretation

a) Suppose that $X_1, \ldots, X_n$ form a random sample from the normal distribution with mean $\mu = 5$ and variance $\sigma^2 = 2.5$.

b) Find a 90 percent confidence interval for the mean (we know the mean!).

c) Generate a sample of size 26 from the sampling distribution and compute the observed 90 percent confidence interval for the mean. Check if μ is in the computed interval or not.

d) Repeat c)

e) Repeat c) 100 times and each time check if μ is in the computed interval or not.