Classical and Bayesian inference

AMS 132

February 16, 2018

Definition of the Distributions

- Here we will discuss another family of distributions: the t (or t Student) distributions.
- These are very important in problems of statistical inference

Example

Suppose that $X_1, \ldots, X_n$ form a random sample from the normal distribution with mean μ and variance σ^2, and that we are interested in how far the sample mean, $\overline{X}_n$ is from the mean μ.
We know that $\sqrt{n}(\overline{X}_n - \mu)/\sigma$ has a standard normal distribution, but we don't know σ.
If we replace σ by an estimator $\widehat{\sigma}$ such as the M.L.E., what is the distribution of $\sqrt{n}(\overline{X}_n - \mu)/\widehat{\sigma}$?

Definition of the Distributions

Definition (t Distributions)

Consider two independent random variables Y and Z, such that Y has the χ^2 distribution with m degrees of freedom and Z has the standard normal distribution. Suppose that a random variable X is defined by the equation

$$X = \frac{Z}{\left(\frac{Y}{m}\right)^{1/2}}.$$

Then the distribution of X is called the t distribution with m degrees of freedom, $m > 0$.

Definition of the Distributions

Theorem (Probability density function)

The p.d.f. of the t distribution with m degrees of freedom is

$$\frac{\Gamma(\frac{m+1}{2})}{(m\pi)^{1/2}\Gamma(\frac{m}{2})}\left(1+\frac{x^2}{m}\right)^{-(m+1)/2}, \quad -\infty < x < \infty.$$

- If X has the t distribution with m degrees of freedom, $E(|X|^k) < \infty$, if $k < m$.
- If X has the t distribution with m degrees of freedom, $E(X) = 0$, if $m > 1$, and $Var(X) = m/(m-2)$, if $m > 2$.
- See plots of the density function in R.

Relation to Random Samples from a Normal Distribution

Theorem

Suppose that $X_1, \ldots, X_n$ form a random sample from the normal distribution with mean μ and variance σ^2. Let $\overline{X}_n$ denote the sample mean, and define

$$\sigma' = \left[\frac{\sum_{i=1}^{n}(X_i - \overline{X}_n)^2}{n-1}\right]^{1/2}.$$

Then $\sqrt{n}(\overline{X}_n - \mu)/\sigma'$ has the t distribution with $n-1$ degrees of freedom.

Proof:

- Note that $\sqrt{n}(\overline{X}_n - \mu)/\sigma'$ and its distribution do not depend on σ!
- Notice that $\sigma' = \left(\frac{n}{n-1}\right)^{1/2}\widehat{\sigma}$, and for large values of n both estimators will be very close.

Relation to Random Samples from 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.

Find the value of c such that $P(U < c) = 0.9$, where $U = \frac{\sqrt{n}(\overline{X}_n - \mu)}{\sigma'}$. What can you say about the difference between the sample mean and the mean?