Hypothesis tests for the 1-sample mean

(Math111-01, Raritan Valley Community College, Summer II 2017)

Notation

- μ is the population mean and $\bar{x} = \frac{\sum x_i}{n}$ is the sample mean (sample size n).
- σ is the population standard deviation and $s = \sqrt{\frac{1}{N-1}\sum (x_i - \bar{x})^2}$ is the sample standard deviation
- H_0 is the null hypothesis (in this chapter, H_0: $\mu = \mu_0$, where μ_0 is the null value) and H_A is the alternative hypotheses
- Z is a standard normal random variable. T is a t-distributed random variable with $n-1$ degrees of freedom.
- z_α is the left endpoint of the right tail of the $z-$curve with area/probability $= \alpha$. t_α is similar but for a $t-$curve with $n-1$ degrees of freedom.
- α is the significance level.

Assumptions

In the following equations, we assume one of the following:

1. The observations are independent (that is, the sampling has to be truly random). In practice, this requires that the sample size be no greater than ten percent of the population size.
2. The population should be nearly normal or the sample size should be large.

To test for normality, a quantile-quantile plot may be used. The less assumption 2 holds faithfully, the more one should rely on the t-distributions over the standard normal distribution.

If σ is not known (it rarely is because it is a population parameter), then use s in its place.

Confidence Intervals

$1-\alpha$ confidence intervals for $\bar{x}$

$$\bar{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$$
$$\bar{x} \pm t_{\alpha/2}\frac{s}{\sqrt{n}}$$

If n is large, then these intervals are very nearly the same.

Test Statistics

Test statistics for $\bar{x}$:

$$\begin{aligned} z &= \frac{\bar{x}-\mu_0}{\sigma/\sqrt{n}} \\ t &= \frac{\bar{x}-\mu_0}{s/\sqrt{n}}. \end{aligned}$$

Under the staded assumptions, z is distributed as a standard normal random variable (prior to sampling). And t is distributed as a t-distribution with $\text{df} = n-1$.

Hypothesis tets with p-values (preferred method)

Alternative hypothesis	p-value (using z)	p-value (using t)
$\mu > \mu_0$	$P(Z > z)$	$P(T > t)$
$\mu < \mu_0$	$P(Z < z)$	$P(T < t)$
$\mu \neq \mu_0$	$\begin{cases} 2P(Z > z) & z \text{ positive} \\ 2P(Z < z) & z \text{ negative} \end{cases}$	$\begin{cases} 2P(T > t) & t \text{ positive} \\ 2P(T < t) & t \text{ negative} \end{cases}$

Reject the null hypothesis only when $p < \alpha$. Finding the p-value precisely is easy to do if using z. But using the tables, one can only find a range of possible p-values using t.

Hypothesis testing with critical values

Alternative hypothesis	rejection region (using z)	rejection region (using t)
$\mu > \mu_0$	$z > z_\alpha$	$t > t_\alpha$
$\mu < \mu_0$	$z < -z_\alpha$	$t < -t_\alpha$
$\mu \neq \mu_0$	$z > z_{\alpha/2}$ or $z < -z_{\alpha/2}$	$t > t_{\alpha/2}$ or $t < -t_{\alpha/2}$