Basic Statistics and Hypothesis Testing

Statistics terminology

Probability density function (*pdf*): A function that plots the probability of observing a particular value. Given a normal distribution, for example, the probability of observing a value close to the mean is much higher than one that is far from the mean.

- Sample: A collection of observations from a population
- Mean: The average value.
- Median: The midpoint of the *pdf*. This is the "midpoint" of the data – half of the values fall on one side of the median, the other half fall on the other side.
- Standard deviation: This measures how widely dispersed the data are – how wide the *pdf* falls around the mean. The variance is the squared-value of the standard deviation.

Hypothesis Testing

The first step in conducting a hypothesis test is to identify the test. This means constructing a null hypothesis (H_o) and an alternative hypothesis (H_a). For example, if you wanted to test whether a sample mean, μ, equals a particular value μ_o, the test would be written as:

$$H_o : \mu = \mu_o$$
$$H_a : \mu \neq \mu_o$$

Now, to test your hypothesis, you need to construct a test statistic. This allows you to do the test using a standard distribution (like a normal distribution, or a student's T distribution). The one you will use most often is the t-statistic:

$$t\ stat = \frac{\mu - \mu_o}{s_\mu}$$

Where s_μ is the standard error. In this case, the standard error would be the estimated standard deviation σ, divided by the square root of the sample size n:

$$s_\mu = \frac{\sigma}{\sqrt{n}}$$

Now, you need to identify the critical value of the t-statistic that will allow you to reject your null hypothesis. The critical value you look for depends on how confident you want to be in your test statistic. A critical value is associated with a confidence level - the most commonly used confidence level is 95%, although 90% and 99% are also used. Other factors, like the sample size, determine which critical value you use to conduct your test. The relationship between confidence level, t-statistic, and the probability density function is explained in more detail on the following page.

In the statements below, α below refers to the significance level. Commonly used significance levels are 1%, 5%, and 10% - these are associated with the aforementioned confidence levels.

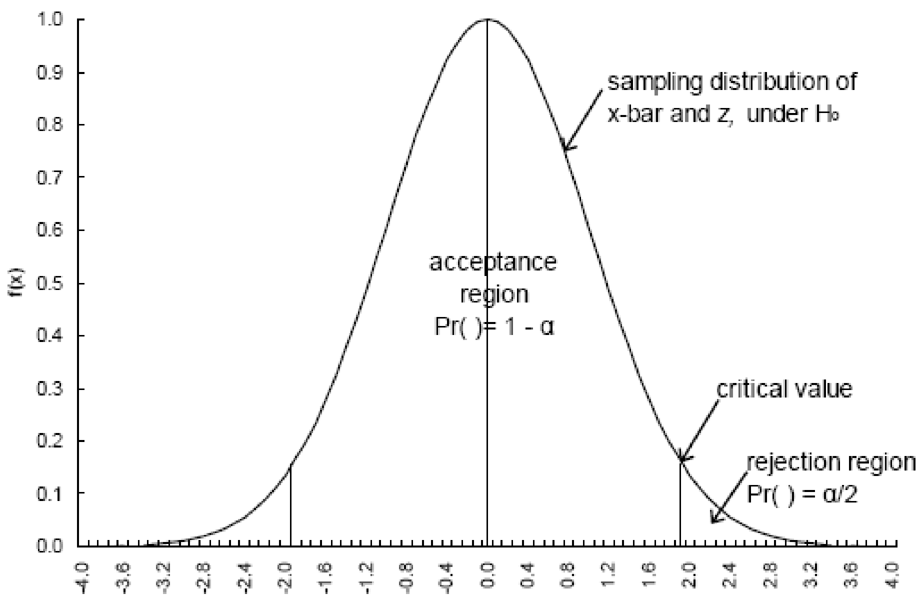
For a two-sided test:

$H_o : \mu = \mu_o$	If $\lvert t\ stat \rvert > t_{\alpha/2,(n-2)} \Rightarrow$ reject H_o	If $\lvert t\ stat \rvert \leq t_{\alpha/2,(n-2)} \Rightarrow$ fail to reject H_o
$H_a : \mu \neq \mu_o$	If $p - value < \alpha \Rightarrow$ reject H_o	If $p - value \geq \alpha \Rightarrow$ fail to reject H_o

For one-sided tests:

$H_o : \mu \geq \mu_o$	If $t\ stat < -t_{\alpha,(n-2)} \Rightarrow$ reject H_o	If $t\ stat \geq -t_{\alpha,(n-2)} \Rightarrow$ fail to reject H_o
$H_a : \mu < \mu_o$	If $^{p-value}/_2 < \alpha$ and $t\ stat < 0 \Rightarrow$ reject H_o	If $^{p-value}/_2 \geq \alpha$ or $t\ stat \geq 0 \Rightarrow$ fail to reject H_o
$H_o : \mu \leq \mu_o$	If $t\ stat > t_{\alpha,(n-2)} \Rightarrow$ reject H_o	If $t\ stat \leq t_{\alpha,(n-2)} \Rightarrow$ fail to reject H_o
$H_a : \mu > \mu_o$	If $^{p-value}/_2 < \alpha$ and $t\ stat > 0 \Rightarrow$ reject H_o	If $^{p-value}/_2 \geq \alpha$ or $t\ stat \leq 0 \Rightarrow$ fail to reject H_o

HYPOTHESIS TEST EXAMPLE: TEST OF $\mu=\mu_0$



$$z = \left(\frac{\bar{x} - \mu_0}{\sigma / \sqrt{n}} \right)$$

Hypothesis:
- H_0: $\mu = \mu_0$
- H_1: $\mu \neq \mu_0$

Significance Level (Probability of Type I Error):
$\alpha = 0.05$

Under H_0, the sample mean $\bar{x}$ has a sampling distribution with mean μ_0 and standard deviation $\sigma/\sqrt{n}$. If the underlying random variable x is normally distributed, or if our sample size is large, then our test statistic z (the standardized version of $\bar{x}$) is distributed standard normal.[1]

If H_0 is true, then there is a 95 percent $(1 - \alpha)$ probability that our estimate of the test statistic z will fall within the acceptance region shown above [-1.96, 1.96]. Thus, we *do not reject* H_0 if z falls within this region ("z is not statistically different from zero"), or *reject* H_0 if it does not.

Equivalently, if H_0 is true, $\bar{x}$ should fall in the region [$\mu_0 - 1.96 * \sigma/\sqrt{n}$, $\mu_0 + 1.96 * \sigma/\sqrt{n}$]. If our estimate of $\bar{x}$ does not, we can reject H_0. Doing a little math,

$$\Pr\left(-1.96 \leq \left(\frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} \right) \leq 1.96 \right) = 0.95 \quad \text{is equivalent to}$$

$$\Pr\left(\mu_0 - \frac{1.96\sigma}{\sqrt{n}} \leq \bar{x} \leq \mu_0 + \frac{1.96\sigma}{\sqrt{n}} \right) = 0.95 \quad \text{and} \qquad \Pr\left(\bar{x} - \frac{1.96\sigma}{\sqrt{n}} \leq \mu_0 \leq \bar{x} + \frac{1.96\sigma}{\sqrt{n}} \right) = 0.95$$

The latter says that—if H_0 is true, μ_0 must fall in a 95 percent confidence interval around $\bar{x}$.

---

[1] If the variance of $\bar{x}$ $\left(\sigma^2/n\right)$ is *not* known then one estimates the standard deviation with the standard *error* se($\bar{x}$)=$s/\sqrt{n}$, and uses the test statistic $t = \left((\bar{x} - \mu_0)/se(\bar{x})\right)$ which follows a *t-distribution* with *n-1* degrees of freedom (unless the sample size is large, then it is approximately standard normal).