Review of One and Two Sample Tests

One Sample Tests: Normality

Assume that the sample of n observations is from a normal population with mean μ and variance σ^2 (abbreviated $N(\mu, \sigma^2)$). Tests of one or sided hypotheses count the number of standard errors that separate the sample mean $\overline{Y}$ from the null hypothesis.

If σ^2 is known, then the standard error of $\overline{Y}$ is $\sigma/\sqrt{n}$ and use z_α values from the normal table. Otherwise, estimate the standard error using the sample variance as $s/\sqrt{n}$ and use values $t_{\alpha,n-1}$ from the t-table with $n-1$ df. The case of known variance is mostly of conceptual interest, though it does make it possible to answer questions of power and sample size (which could not be easily done from the t-table).

For testing the hypotheses

$$\begin{aligned} H_0 &: \mu \le \mu_0 \\ H_a &: \mu > \mu_0 \end{aligned}$$

use the test statistic (z on p 314, t on p 325)

$$z = \frac{\overline{Y} - \mu_0}{\sigma/\sqrt{n}} \quad \text{or} \quad t = \frac{\overline{Y} - \mu_0}{s/\sqrt{n}} \ .$$

If the value of z (or t) is negative, $\overline{Y}$ lies in the region specified by H_0 so there's no reason to reject. If the value of z (or t) is positive, then reject H_0 if

$$z > z_\alpha \quad \text{or} \quad t > t_{\alpha,n-1} \ .$$

Equivalently, reject H_0 if the *p-value* is less than α (see §8.3, p 321) or (for two-sided tests) if the 100(1-α)% confidence interval does *not* include μ_0 (§8.7, p 335).

Power and Sample Size

To find the power of a test (or β which is 1 minus the power), you have to figure out the probability of rejecting H_0 when it is false. You can do this only when σ^2 is known. For the one-sided hypotheses given above, you can get the power at some $\mu_a > \mu_0$ as (p 319)

$$\begin{aligned} P\{\text{reject } H_0\} = 1 - \beta &= P\{\overline{Y} > \mu_0 + z_\alpha \frac{\sigma}{\sqrt{n}} | \mu = \mu_a\} \\ &= P\{\frac{\overline{Y} - \mu_a}{\sigma/\sqrt{n}} > z_\alpha + \frac{\mu_0 - \mu_a}{\sigma/\sqrt{n}}\} \\ &= P\{N(0,1) > z_\alpha + \frac{\mu_0 - \mu_a}{\sigma/\sqrt{n}}\} \end{aligned}$$

Plug in values for the terms on the right and look up the value in the normal table to find 1-β.

To find the sample size n required to obtain chosen values for α and β at some μ_a, there's a similar formula. The last equation given above implies that

$$\beta = P\{N(0,1) < \underbrace{z_\alpha + \frac{\mu_0 - \mu_a}{\sigma/\sqrt{n}}}_{-z_\beta}\}$$

or

$$-z_\beta = z_\alpha + \frac{\mu_0 - \mu_a}{\sigma/\sqrt{n}} \ .$$

Solving for the required sample size n gives the formula seen on old exams,

$$n = \frac{\sigma^2 (z_\alpha + z_\beta)^2}{(\mu_a - \mu_0)^2} \ .$$

Two-sample Tests

The big changes from one-sample tests to two-sample tests are

- H_0 hypothesizes values for $\mu_1 - \mu_2$, often $H_0 : \mu_1 - \mu_2 = 0$.
- The needed standard error is $SE(\overline{Y}_1 - \overline{Y}_2)$.
- You have to decide whether the samples are dependent or independent.

When the samples are dependent (in particular, have noticeable correlation), you should use the *paired t-test* which allows for the correlation (§9.3, p 370). The paired t-test is just a one-sample test based on the differences. If the two samples are independent (no noticeable correlation), then use the *two-sample t-test* (§9.1, p 351). Both tests resemble one-sample tests in that they count the number of standard errors separating $\overline{Y}_1 - \overline{Y}_2$ from the values specified by H_0 (usually 0).

An important variant of the two-sample test (p 357) allows for the two samples to have different variances. Use this one if the two sample standard deviations are differ by much (e.g., one is twice the size of the other).

Tests for Proportions

For a test of $H_0 : \pi = \pi_0$, use the normal z test formed as (p 332)

$$z = \frac{\hat{\pi} - \pi_0}{\sqrt{\pi_0(1-\pi_0)/n}}$$

as long as $n\pi_0$ and $n(1-\pi_0)$ are about 5 or larger. For a two sample test of $H_0 : \pi_1 - \pi_2 = D_0$, use ($D_0 = 0$ usually, p 395)

$$z = \frac{(\hat{\pi}_1 - \hat{\pi}_2) - D_0}{\sqrt{\frac{\hat{\pi}_1(1-\hat{\pi}_1)}{n_1} + \frac{\hat{\pi}_2(1-\hat{\pi}_2)}{n_2}}}$$

Nonparametric Tests

These tests avoid the assumption of normality and accommodate outliers, but require the other assumptions (such as independent observations). The simplest is the *sign test* discussed in class (p 374). Used in the context of a paired t-test, it works like a binomial test for the number of heads in n tosses of a fair coin.

Better nonparametric methods (ie, ones that work almost as well as the t-tests when the data is normal) use ranks. In one-sample problems, the *signed-rank* test for $H_0 : \mu = 0$ uses a z-score from (p 375)

$$T_+ = \sum \text{ranks of positive values}, \quad z = \frac{T_+ - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}}$$

to see if the ranks of the positive values differ from what is expected under H_0. Since paired comparisons *reduce to* one-sample tests, the signed-rank test is an alternative to the paired t-test.

For two-sample problems, you can use the *rank-sum test.* Its test statistic is found by merging the data from both samples, ordering them, and then counting the ranks of one of the two samples. The test again works by forming a z score that compares the observed sum T to what is expected under H_0 (p 365)

$$T = \sum \text{ ranks of one group }, \quad z = \frac{T - \frac{n_1(n_1+n_2+1)}{2}}{\sqrt{\frac{n_1 n_2}{12}(n_1+n_2+1)}}$$