W09

Population propotion

Large sample

Let $\hat{p}_1 = X/m$ and $\hat{p}_2 = Y/n$, where $X \sim \text{Bin}(m, p_1)$ and $Y \sim \text{Bin}(n, p_2)$ with X and Y independent variables. Then

$$E(\hat{p}_1 - \hat{p}_2) = p_1 - p_2$$

so $\hat{p}_1 - \hat{p}_2$ is an unbiased estimator of $p_1 - p_2$, and

$$V(\hat{p}_1 - \hat{p}_2) = \frac{p_1 q_1}{m} + \frac{p_2 q_2}{n} \quad (\text{where } q_i = 1 - p_i) \tag{9.3}$$

Null hypothesis: $H_0: p_1 - p_2 = 0$

Test statistic value (large samples): $z = \dfrac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}\hat{q}\left(\frac{1}{m} + \frac{1}{n}\right)}}$

Alternative Hypothesis	Rejection Region for Approximate Level α Test
$H_a: p_1 - p_2 > 0$	$z \ge z_\alpha$
$H_a: p_1 - p_2 < 0$	$z \le -z_\alpha$
$H_a: p_1 - p_2 \ne 0$	either $z \ge z_{\alpha/2}$ or $z \le -z_{\alpha/2}$

A P-value is calculated in the same way as for previous z tests.
The test can safely be used as long as $m\hat{p}_1$, $m\hat{q}_1$, $n\hat{p}_2$, and $n\hat{q}_2$ are all at least 10.

Type II and sample size

$$\sigma_{\hat{p}_1 - \hat{p}_2} = \sqrt{\frac{p_1 q_1}{m} + \frac{p_2 q_2}{n}}$$

Alternative Hypothesis	$\beta(p_1, p_2)$
$H_a: p_1 - p_2 > 0$	$\Phi\left[\dfrac{z_\alpha\sqrt{\bar{p}\,\bar{q}\left(\frac{1}{m} + \frac{1}{n}\right)} - (p_1 - p_2)}{\sigma}\right]$
$H_a: p_1 - p_2 < 0$	$1 - \Phi\left[\dfrac{-z_\alpha\sqrt{\bar{p}\,\bar{q}\left(\frac{1}{m} + \frac{1}{n}\right)} - (p_1 - p_2)}{\sigma}\right]$

Alternative Hypothesis	$\beta(p_1, p_2)$
$H_a: p_1 - p_2 \ne 0$	$\Phi\left[\dfrac{z_{\alpha/2}\sqrt{\bar{p}\,\bar{q}\left(\frac{1}{m} + \frac{1}{n}\right)} - (p_1 - p_2)}{\sigma}\right] - \Phi\left[\dfrac{-z_{\alpha/2}\sqrt{\bar{p}\,\bar{q}\left(\frac{1}{m} + \frac{1}{n}\right)} - (p_1 - p_2)}{\sigma}\right]$

where $\bar{p} = (mp_1 + np_2)/(m + n)$, $\bar{q} = (mq_1 + nq_2)/(m + n)$, and σ is given by (9.6).

For the case $m = n$, the level α test has type II error probability β at the alternative values p_1, p_2 with $p_1 - p_2 = d$ when

$$n = \frac{\left[z_\alpha\sqrt{(p_1 + p_2)(q_1 + q_2)/2} + z_\beta\sqrt{p_1 q_1 + p_2 q_2}\right]^2}{d^2} \tag{9.7}$$

for an upper- or lower-tailed test, with $\alpha/2$ replacing α for a two-tailed test.

Large sample size

A CI for $p_1 - p_2$ with confidence level approximately $100(1 - \alpha)\%$ is

$$\hat{p}_1 - \hat{p}_2 \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1\hat{q}_1}{m} + \frac{\hat{p}_2\hat{q}_2}{n}}$$

This interval can safely be used as long as $m\hat{p}_1$, $m\hat{q}_1$, $n\hat{p}_2$, and $n\hat{q}_2$ are all at least 10.

CI $\hat{p}_1 - \hat{p}_2 \pm z_{\alpha/2}\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{m} + \frac{1}{n}\right)}$

Population variances

F distribution

$$F = \frac{X_1/v_1}{X_2/v_2}$$

$$F_{1-\alpha, v_1, v_2} = 1/F_{\alpha, v_2, v_1}.$$

Let $X_1, \ldots, X_m$ be a random sample from a normal distribution with variance σ_1^2, let $Y_1, \ldots, Y_n$ be another random sample (independent of the X_i's) from a normal distribution with variance σ_2^2, and let S_1^2 and S_2^2 denote the two sample variances. Then the rv

$$F = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \tag{9.9}$$

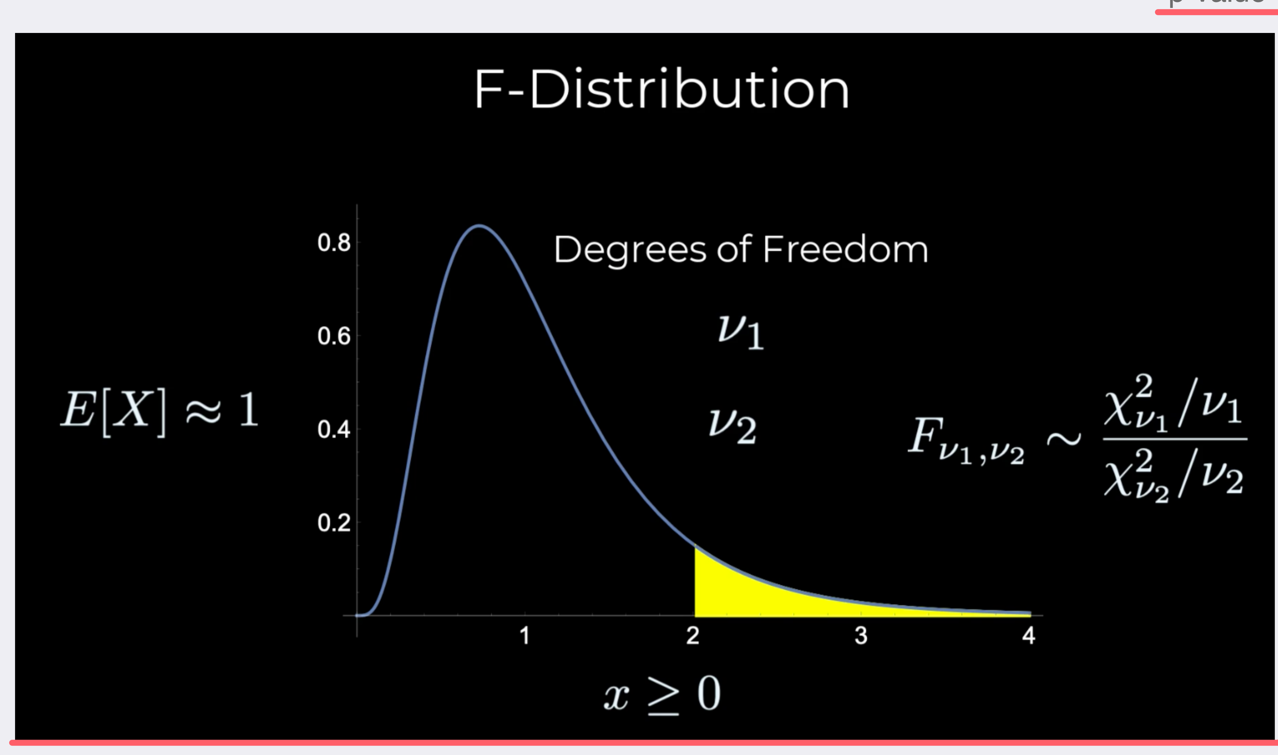
has an F distribution with $v_1 = m - 1$ and $v_2 = n - 1$.

Null hypothesis: $H_0: \sigma_1^2 = \sigma_2^2$
Test statistic value: $f = s_1^2/s_2^2$

Alternative Hypothesis	Rejection Region for a Level α Test
$H_a: \sigma_1^2 > \sigma_2^2$	$f \ge F_{\alpha, m-1, n-1}$
$H_a: \sigma_1^2 < \sigma_2^2$	$f \le F_{1-\alpha, m-1, n-1}$
$H_a: \sigma_1^2 \ne \sigma_2^2$	either $f \ge F_{\alpha/2, m-1, n-1}$ or $f \le F_{1-\alpha/2, m-1, n-1}$

Since critical values are tabled only for $\alpha = .10, .05, .01$, and $.001$, the two-tailed test can be performed only at levels $.20, .10, .02$, and $.002$. Other F critical values can be obtained from statistical software.

p-value



Excel

F stat: FINV(pvalue, df1, df2)

p-value: FDIST(fstat, df1, df2)

Two Sample, Population means: z Test and CI

Normal population with known sigma

The expected value of $\bar{X} - \bar{Y}$ is $\mu_1 - \mu_2$, so $\bar{X} - \bar{Y}$ is an unbiased estimator of $\mu_1 - \mu_2$. The standard deviation of $\bar{X} - \bar{Y}$ is

$$\sigma_{\bar{X}-\bar{Y}} = \sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}$$

Null hypothesis: $H_0: \mu_1 - \mu_2 = \Delta_0$

Test statistic value: $z = \dfrac{\bar{x} - \bar{y} - \Delta_0}{\sqrt{\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n}}}$

Alternative Hypothesis	Rejection Region for Level α Test
$H_a: \mu_1 - \mu_2 > \Delta_0$	$z \ge z_\alpha$ (upper-tailed)
$H_a: \mu_1 - \mu_2 < \Delta_0$	$z \le -z_\alpha$ (lower-tailed)
$H_a: \mu_1 - \mu_2 \ne \Delta_0$	either $z \ge z_{\alpha/2}$ or $z \le -z_{\alpha/2}$ (two-tailed)

Because these are z tests, a P-value is computed as it was for the z tests in Chapter 8 [e.g., P-value $= 1 - \Phi(z)$ for an upper-tailed test].

Alternative Hypothesis	$\beta(\Delta') = P(\text{type II error when } \mu_1 - \mu_2 = \Delta')$
$H_a: \mu_1 - \mu_2 > \Delta_0$	$\Phi\left(z_\alpha - \dfrac{\Delta' - \Delta_0}{\sigma}\right)$
$H_a: \mu_1 - \mu_2 < \Delta_0$	$1 - \Phi\left(-z_\alpha - \dfrac{\Delta' - \Delta_0}{\sigma}\right)$
$H_a: \mu_1 - \mu_2 \ne \Delta_0$	$\Phi\left(z_{\alpha/2} - \dfrac{\Delta' - \Delta_0}{\sigma}\right) - \Phi\left(-z_{\alpha/2} - \dfrac{\Delta' - \Delta_0}{\sigma}\right)$

where $\sigma = \sigma_{\bar{X}-\bar{Y}} = \sqrt{(\sigma_1^2/m) + (\sigma_2^2/n)}$

Beta and sample size

the power is 1 - beta

$$\frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n} = \frac{(\Delta' - \Delta_0)^2}{(z_\alpha + z_\beta)^2}$$

$$m = n = \frac{(\sigma_1^2 + \sigma_2^2)(z_\alpha + z_\beta)^2}{(\Delta' - \Delta_0)^2}$$

Large sample test

Use of the test statistic value

$$z = \frac{\bar{x} - \bar{y} - \Delta_0}{\sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}}$$

along with the previously stated upper-, lower-, and two-tailed rejection regions based on z critical values gives large-sample tests whose significance levels are approximately α. These tests are usually appropriate if both $m > 40$ and $n > 40$. A P-value is computed exactly as it was for our earlier z tests.

Provided that m and n are both large, a CI for $\mu_1 - \mu_2$ with a confidence level of approximately $100(1 - \alpha)\%$ is

$$\bar{x} - \bar{y} \pm z_{\alpha/2}\sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}$$

where $-$ gives the lower limit and $+$ the upper limit of the interval. An upper or a lower confidence bound can also be calculated by retaining the appropriate sign ($+$ or $-$) and replacing $z_{\alpha/2}$ by z_α.

$$n = \frac{4z_{\alpha/2}^2(\sigma_1^2 + \sigma_2^2)}{w^2}$$

Round up

t-Test and CI

When the population distributions are both normal, the standardized variable

$$T = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{m} + \frac{S_2^2}{n}}} \tag{9.2}$$

has approximately a t distribution with df v estimated from the data by

$$v = \frac{\left(\frac{s_1^2}{m} + \frac{s_2^2}{n}\right)^2}{\frac{(s_1^2/m)^2}{m-1} + \frac{(s_2^2/n)^2}{n-1}} = \frac{[(se_1)^2 + (se_2)^2]^2}{\frac{(se_1)^4}{m-1} + \frac{(se_2)^4}{n-1}}$$

where

$$se_1 = \frac{s_1}{\sqrt{m}}, \quad se_2 = \frac{s_2}{\sqrt{n}}$$

(round v down to the nearest integer).

The **two-sample t confidence interval for $\mu_1 - \mu_2$** with confidence level $100(1 - \alpha)\%$ is then

$$\bar{x} - \bar{y} \pm t_{\alpha/2, v}\sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}$$

A one-sided confidence bound can be calculated as described earlier.
The **two-sample t test** for testing $H_0: \mu_1 - \mu_2 = \Delta_0$ is as follows:

Test statistic value: $t = \dfrac{\bar{x} - \bar{y} - \Delta_0}{\sqrt{\frac{s_1^2}{m} + \frac{s_2^2}{n}}}$

Alternative Hypothesis	Rejection Region for Approximate Level α Test
$H_a: \mu_1 - \mu_2 > \Delta_0$	$t \ge t_{\alpha, v}$ (upper-tailed)
$H_a: \mu_1 - \mu_2 < \Delta_0$	$t \le -t_{\alpha, v}$ (lower-tailed)
$H_a: \mu_1 - \mu_2 \ne \Delta_0$	either $t \ge t_{\alpha/2, v}$ or $t \le -t_{\alpha/2, v}$ (two-tailed)

A P-value can be computed as described in Section 8.4 for the one-sample t test.

Pooled t procedures $S_p^2 = \dfrac{m-1}{m+n-2} \cdot S_1^2 + \dfrac{n-1}{m+n-2} \cdot S_2^2$ normal, equal sigma

Type II error probabilities P361

Recall CLT $\bar{X} \sim N\left(\mu_{\bar{x}} = \mu_x,\ \sigma_{\bar{x}} = \frac{\sigma_x}{\sqrt{n}}\right)$

Paired data

The Paired t Test

Null hypothesis: $H_0: \mu_D = \Delta_0$ (where $D = X - Y$ is the difference between the first and second observations within a pair, and $\mu_D = \mu_1 - \mu_2$)

Test statistic value: $t = \dfrac{\bar{d} - \Delta_0}{s_D/\sqrt{n}}$ (where $\bar{d}$ and s_D are the sample mean and standard deviation, respectively, of the d_i's)

Alternative Hypothesis	Rejection Region for Level α Test
$H_a: \mu_D > \Delta_0$	$t \ge t_{\alpha, n-1}$
$H_a: \mu_D < \Delta_0$	$t \le -t_{\alpha, n-1}$
$H_a: \mu_D \ne \Delta_0$	either $t \ge t_{\alpha/2, n-1}$ or $t \le -t_{\alpha/2, n-1}$

A P-value can be calculated as was done for earlier t tests.

The paired t CI for μ_D is

$$\bar{d} \pm t_{\alpha/2, n-1} \cdot s_D/\sqrt{n}$$

A one-sided confidence bound results from retaining the relevant sign and replacing $t_{\alpha/2}$ by t_α.

Paired data and two samples t procedures

$$V(X \pm Y) = V(X) + V(Y) \pm 2\,\text{Cov}(X, Y)$$

$$\rho = \text{Corr}(X, Y) = \text{Cov}(X, Y)/[\sqrt{V(X)} \cdot \sqrt{V(Y)}]$$

$$V(\bar{X} - \bar{Y}) = V(\bar{D}) = V\left(\frac{1}{n}\Sigma D_i\right) = \frac{V(D_i)}{n} = \frac{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}{n}$$

1. If there is great heterogeneity between experimental units and a large correlation within experimental units (large positive ρ), then the loss in degrees of freedom will be compensated for by the increased precision associated with pairing, so a paired experiment is preferable to an independent-samples experiment.
2. If the experimental units are relatively homogeneous and the correlation within pairs is not large, the gain in precision due to pairing will be outweighed by the decrease in degrees of freedom, so an independent-samples experiment should be used.