

Parametric Statistics

Hypothesis Testing

Sofia Triantafyllou

sof.triantafyllou@gmail.com

University of Crete
Department of Mathematics and Applied Mathematics

January 7, 2024

Lecture Summary

9.5 The t-Test

9.6 Comparing the Means of Two Normal Distributions

9.7 The F-Distributions

The t -Test

- ▶ The t -Test is a test for hypotheses concerning the mean parameter in the normal distribution when the variance is also unknown.
- ▶ The test is based on the t distribution

The setup for the next few slides:

- ▶ Let X_1, \dots, X_n be i.i.d. $N(\mu, \sigma^2)$ and consider the hypotheses

$$H_0 : \mu \leq \mu_0 \quad \text{vs.} \quad H_1 : \mu > \mu_0$$

The parameter space here is $-\infty < \mu < \infty$ and $\sigma^2 > 0$, i.e.

$$\Omega = (-\infty, \infty) \times (0, \infty)$$

And

$$\Omega_0 = (-\infty, \mu_0] \times (0, \infty) \quad \text{and} \quad \Omega_1 = (\mu_0, \infty) \times (0, \infty)$$

The one-sided t -Test

- ▶ Let

$$U = \frac{\sqrt{n}(\bar{X}_n - \mu_0)}{\sigma'} \quad \text{where } \sigma' = \left(\frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X}_n)^2 \right)^{1/2}$$

- ▶ If $\mu = \mu_0$ then U has the t distribution
- ▶ Tests based on U are called t tests

The one-sided t -Test

- ▶ Let T_n^{-1} be the quantile function of the t_n distribution

Theorem (Theorem 9.5.1)

The test δ that rejects H_0 if $U \geq T_{n-1}^{-1}(1 - \alpha_0)$ has size α_0 and a power function with the following properties

- (i) $\pi(\mu_0, \sigma^2 | \delta) = \alpha_0$
- (ii) $\pi(\mu, \sigma^2 | \delta) < \alpha_0$ for $\mu < \mu_0$
- (iii) $\pi(\mu, \sigma^2 | \delta) > \alpha_0$ for $\mu > \mu_0$
- (iv) $\pi(\mu, \sigma^2 | \delta) \rightarrow 0$ as $\mu \rightarrow -\infty$
- (v) $\pi(\mu, \sigma^2 | \delta) \rightarrow 1$ as $\mu \rightarrow \infty$

The complete power function

To calculate the power function $\pi(\mu, \sigma^2 \mid \delta)$ exactly we need the non-central t_m distributions:

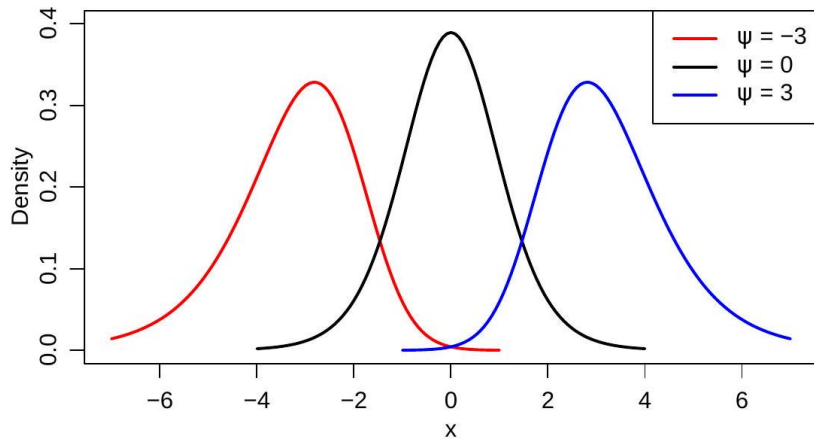
Definition

Let $W \sim N(\psi, 1)$ and $Y \sim \chi_m^2$ be independent. The distribution of

$$X = \frac{W}{(Y/m)^{1/2}}$$

is called the non-central t distribution with m degrees of freedom and non-centrality parameter ψ

Non-central t_m distribution



The complete power function

Theorem (Theorem 9.5.3)

U has the non-central t_{n-1} distribution with non-centrality parameter $\psi = \sqrt{n}(\mu - \mu_0) / \sigma$

The power function of the t-test that rejects H_0 if

$U \geq T_{n-1}^{-1}(1 - \alpha_0) = c_1$ is

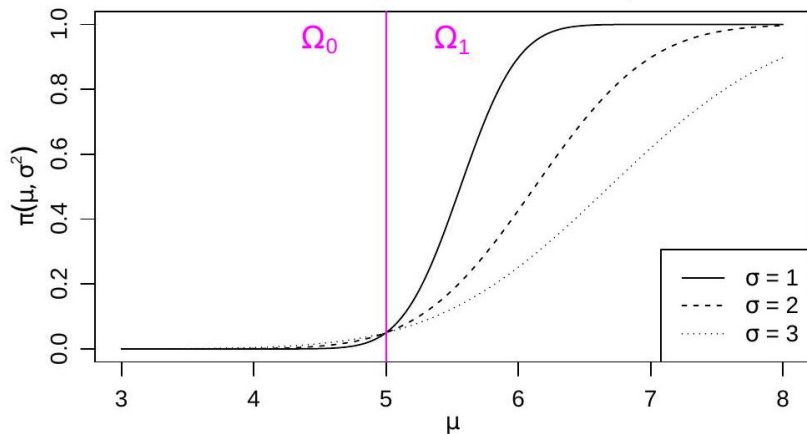
$$\pi(\mu, \sigma^2 | \delta) = 1 - T_{n-1}(c_1 | \psi)$$

- ▶ Can use the R function $1 - \text{pt}(q = c_1, \text{df} = n - 1, \text{ncp} = \text{psi})$

Power function for the one-sided t -test

Example: $n = 10, \mu_0 = 5, \alpha_0 = 0.05$

Power function for the size 0.05 t -test, with $n = 10$



Note that the power function is a function of both σ^2 and μ

p-value for the one-sided t -Test

Theorem (9.5.2:p-values for t tests)

Let $H_0 : \mu \leq \mu_0$ vs. $H_1 : \mu > \mu_0$. Let δ be the one sided t -test for H_0, H_1 . Let u be the observed value of U . The p -value is $1 - T_{n-1}(u)$.

Example: Acid Concentration in Cheese (Example 8.5.4)

- ▶ Have a random sample of $n = 10$ lactic acid measurements from cheese, assumed to be from a normal distribution with unknown mean and variance.
- ▶ Observed: $\bar{x}_n = 1.379$ and $\sigma' = 0.3277$
- ▶ Perform the level $\alpha_0 = 0.05$ t -test of the hypotheses

$$H_0 : \mu \leq 1.2 \quad \text{vs} \quad H_1 : \mu > 1.2$$

- ▶ Compute the p-value

The other one-sided t -Test

- ▶ Now consider the hypotheses

$$H_0 : \mu \geq \mu_0 \quad \text{vS.} \quad H_1 : \mu < \mu_0$$

Corollary (9.5.1)

The test δ that rejects H_0 if $U \leq T_{n-1}^{-1}(\alpha_0)$ has size α_0 and a power function with the following properties

- (i) $\pi(\mu_0, \sigma^2 | \delta) = \alpha_0$
- (ii) $\pi(\mu, \sigma^2 | \delta) > \alpha_0$ for $\mu < \mu_0$
- (iii) $\pi(\mu, \sigma^2 | \delta) < \alpha_0$ for $\mu > \mu_0$
- (iv) $\pi(\mu, \sigma^2 | \delta) \rightarrow 1$ as $\mu \rightarrow -\infty$
- (v) $\pi(\mu, \sigma^2 | \delta) \rightarrow 0$ as $\mu \rightarrow \infty$

Power function and p-value for the other one-sided t -Test

Theorem (9.5.2: p-values for t Tests)

Let u be the observed value of U . The p-value is $T_{n-1}(u)$.

Theorem (Theorem 9.5.3)

U has the non-central t_{n-1} distribution with non-centrality parameter $\psi = \sqrt{n}(\mu - \mu_0)/\sigma$.

The power function of the t-test that rejects H_0 in (2) if $U \leq T_{n-1}^{-1}(\alpha_0) = c_2$ is

$$\pi(\mu, \sigma^2 \mid \delta) = T_{n-1}(c_2 \mid \psi)$$

Two-sided t -test

Consider now the test with a two-sided alternative hypothesis:

$$H_0 : \mu = \mu_0 \quad \text{vs.} \quad H_1 : \mu \neq \mu_0$$

- ▶ Let δ be the test that rejects H_0 iff
$$|U| \geq T_{n-1}^{-1}(1 - \alpha_0/2) = c$$
- ▶ Then δ is a size α_0 test
- ▶ The power function is

$$\pi(\mu, \sigma^2 | \delta) = T_{n-1}(-c | \psi) + 1 - T_{n-1}(c | \psi)$$

- ▶ If u is the observed value of U then the p-value is
$$2(1 - T_{n-1}(|u|))$$

The t test is a likelihood ratio test (see p. 583 - 585 in the book)

The paired t -test

Sometimes we are measuring the same variable under two different conditions

- ▶ National Transportation Safety Board crash test dummy experiment:
- ▶ For each car, place
 - ▶ One dummy in the driver's seat.
 - ▶ One dummy in the passenger's seat.
 - ▶ Measure the head injuries for each dummy
- ▶ You want to compare: On average, which seat suffers the most injuries.
- ▶ Take X_1, \dots, X_n to be the head injury in the driver's side and Y_1, \dots, Y_n be head injury in the passenger's side.
- ▶ Test if the means of the two distributions are the same.

The two-sample t -test

Comparing the means of two populations

- ▶ X_1, \dots, X_m i.i.d. $N(\mu_1, \sigma^2)$ and
- ▶ Y_1, \dots, Y_n i.i.d. $N(\mu_2, \sigma^2)$
- ▶ The variance is the same for both samples, but unknown

We are interested in testing one of these hypotheses:

- $H_0 : \mu_1 \leq \mu_2$ vs. $H_1 : \mu_1 > \mu_2$
- $H_0 : \mu_1 \geq \mu_2$ vs. $H_1 : \mu_1 < \mu_2$
- $H_0 : \mu_1 = \mu_2$ vs. $H_1 : \mu_1 \neq \mu_2$

Power function is now a function of 3 parameters:

$$\pi(\mu_1, \mu_2, \sigma^2 \mid \delta)$$

Two-sample t statistic

Let $\bar{X}_m = \frac{1}{m} \sum_{i=1}^m x_i$ and $\bar{Y}_n = \frac{1}{n} \sum_{i=1}^n Y_i$

$$S_X^2 = \sum_{i=1}^m (X_i - \bar{X}_m)^2 \quad \text{and} \quad S_Y^2 = \sum_{i=1}^n (Y_i - \bar{Y}_n)^2$$

$$U = \frac{\sqrt{m+n-2} (\bar{X}_m - \bar{Y}_n)}{\left(\frac{1}{m} + \frac{1}{n}\right)^{1/2} (S_X^2 + S_Y^2)^{1/2}}$$

- ▶ Theorem 9.6.1: If $\mu_1 = \mu_2$ then $U \sim t_{m+n-2}$
- ▶ Theorem 9.6.4: For any μ_1 and μ_2 , U has the non-central t_{m+n-2} distribution with non-centrality parameter

$$\psi = \frac{\mu_1 - \mu_2}{\sigma(1/m + 1/n)^{1/2}}$$

Two-sample t test - summary

a) $H_0 : \mu_1 \leq \mu_2$ vs. $H_1 : \mu_1 > \mu_2$

▶ Level α_0 test: Reject H_0 iff $U \geq T_{m+n-2}^{-1}(1 - \alpha_0)$

▶ p-value: $1 - T_{m+n-2}(u)$

b) $H_0 : \mu_1 \geq \mu_2$ vs. $H_1 : \mu_1 < \mu_2$

▶ Level α_0 test: Reject H_0 iff $U \leq T_{m+n-2}^{-1}(\alpha_0)$

▶ p-value: $T_{m+n-2}(u)$

c) $H_0 : \mu_1 = \mu_2$ vs. $H_1 : \mu_1 \neq \mu_2$

▶ Level α_0 test: Reject H_0 iff $|U| \geq T_{m+n-2}^{-1}(1 - \alpha_0/2)$

▶ p-value: $2(1 - T_{m+n-2}(|u|))$

The two-sample t -test is a likelihood ratio test (see p. 592)

Two-sample t test - unequal variances

- ▶ We can extend the two sample t -test to a problem where the variances of the X_i 's and Y_j 's are not equal but the ratio of them is known, i.e. $\sigma_1^2 = k\sigma_2^2$
- ▶ Not very practical

In general, the problem where the variances are not equal is very hard.

- ▶ Proposed test-statistics do not have known distribution, but approximations have been obtained
- ▶ Example: The Welch statistic

$$V = \frac{\bar{X}_m - \bar{Y}_n}{\left(\frac{S_X^2}{m(m-1)} + \frac{S_Y^2}{n(n-1)} \right)^{1/2}}$$

can be approximated by a t distribution

- ▶ Example: The distribution of the likelihood ratio statistic can be approximated by the χ_1^2 distribution if the sample size is large enough

F-distributions

- ▶ In light of the previous slide, it would be nice to have a test of whether the variances in the two normal populations are equal \rightarrow need the $F_{m,n}$ distributions

Definition

$F_{m,n}$ -distributions Let $Y \sim \chi_m^2$ and $W \sim \chi_n^2$ be independent. The distribution of

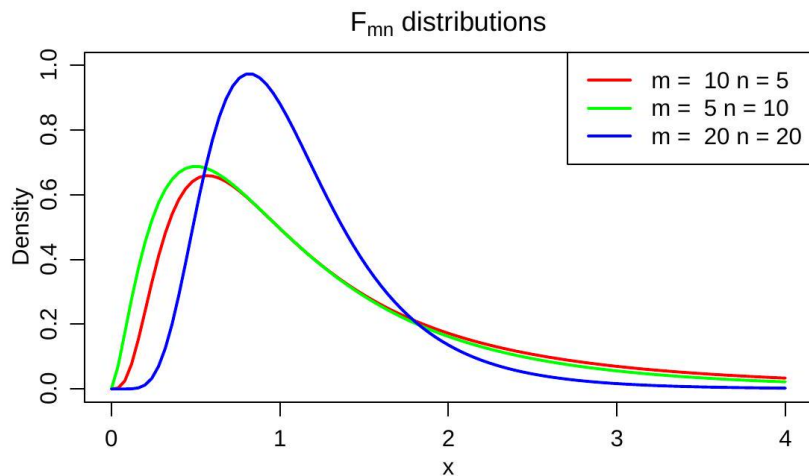
$$X = \frac{Y/m}{W/n} = \frac{nY}{mW}$$

is called the F distribution with m and n degrees of freedom

The pdf of the $F_{m,n}$ distribution is

$$f(x) = \frac{\Gamma((m+n)/2)m^{m/2}n^{n/2}}{\Gamma(m/2)\Gamma(n/2)} \frac{x^{m/2-1}}{(mx+n)^{(m+n)/2}} \quad x > 0$$

F-distributions



- ▶ The 0.95 and 0.975 quantiles of the $F_{m,n}$ distribution is tabulated in the back of the book for a few combinations of m and n

Properties of F distributions

Theorem 9.7.2: Two properties

(i) If $X \sim F_{m,n}$ then $1/X \sim F_{n,m}$

(ii) If $Y \sim t_n$ then $Y^2 \sim F_{1,n}$

Comparing the variances of two populations

- ▶ X_1, \dots, X_m i.i.d. $N(\mu_1, \sigma_1^2)$ and
- ▶ Y_1, \dots, Y_n i.i.d. $N(\mu_2, \sigma_2^2)$
- ▶ All four parameters unknown

Consider the hypotheses:

$$H_0 : \sigma_1^2 \leq \sigma_2^2 \quad \text{vs} \quad H_1 : \sigma_1^2 > \sigma_2^2$$

and the test that rejects H_0 if $V \geq c$, where

$$V = \frac{S_X^2 / (m - 1)}{S_Y^2 / (n - 1)}$$

This test is called an F -test

- ▶ $\frac{\sigma_2^2}{\sigma_1^2} V \sim F_{m-1, n-1}$
- ▶ If $\sigma_1^2 = \sigma_2^2$ then $V \sim F_{m-1, n-1}$

The F test

Let c be the $1 - \alpha_0$ quantile of the F distribution with $m - 1$ and $n - 1$ degrees of freedom, and let $G_{m-1, n-1}$ be the c.d.f. of that F distribution. Let δ be test that rejects H_0 in when $V \geq c$. The power function $\pi(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2 \mid \delta)$ satisfies the following properties:

- ▶ $\pi(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2 \mid \delta) = 1 - G_{m-1, n-1}\left(\frac{\sigma_2^2}{\sigma_1^2}c\right)$,
- ▶ $\pi(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2 \mid \delta) = \alpha_0$ when $\sigma_1^2 = \sigma_2^2$
- ▶ $\pi(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2 \mid \delta) < \alpha_0$ when $\sigma_1^2 < \sigma_2^2$,
- ▶ $\pi(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2 \mid \delta) > \alpha_0$ when $\sigma_1^2 > \sigma_2^2$,
- ▶ $\pi(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2 \mid \delta) \rightarrow 0$ as $\sigma_1^2/\sigma_2^2 \rightarrow 0$,
- ▶ $\pi(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2 \mid \delta) \rightarrow 1$ as $\sigma_1^2/\sigma_2^2 \rightarrow \infty$.

The test δ has level α_0 . The p -value when $V = v$ is observed equals $1 - G_{m-1, n-1}(v)$

Comparing the variances of two populations

- ▶ X_1, \dots, X_m i.i.d. $N(\mu_1, \sigma_1^2)$ and
- ▶ Y_1, \dots, Y_n i.i.d. $N(\mu_2, \sigma_2^2)$
- ▶ All four parameters unknown

Consider the hypotheses:

$$H_0 : \sigma_1^2 = \sigma_2^2 \quad \text{vs} \quad H_1 : \sigma_1^2 \neq \sigma_2^2$$

and the test that rejects H_0 if $V \geq c$, where

$$V = \frac{S_X^2 / (m - 1)}{S_Y^2 / (n - 1)}$$

This test is called an F -test

- ▶ $\frac{\sigma_2^2}{\sigma_1^2} V \sim F_{m-1, n-1}$
- ▶ If $\sigma_1^2 = \sigma_2^2$ then $V \sim F_{m-1, n-1}$

The two-sided F test

If we want to compare:

$$H_0 : \sigma_1^2 = \sigma_2^2 \quad \text{vs.} \quad H_1 : \sigma_1^2 \neq \sigma_2^2$$

- ▶ We will reject H_0 if $V \leq c_1$ or $V \geq c_2$.
- ▶ Typically we choose c_1, c_2 such that $P(V \leq c_1) = P(V \geq c_2) = \alpha_0/2$.
- ▶ Then, p-value $2 * \min\{1 - G_{m-1, n-1}(v), G_{m-1, n-1}(v)\}$, where v is the observed value of V .