Bus 274: Further Statistics For Business

Harald Schmidbauer

İSTANBUL BİLGİ ÜNİVERSİTESİ

About These Slides

- The present slides are not self-contained; they need to be explained and discussed. This will be done in the lectures.

- Even though being a "work in progress" and subject to revision, the slides constitute copyrighted material.
  If you want to reproduce or copy anything from the slides, please ask:

Harald Schmidbauer	**harald** at **hs-stat** dot **com**
Angi Rösch	**angi.r** at **t-online** dot **de**

- The slides were produced using LaTeX and R (the R project; www.R-project.org) on a GNU/Linux system.

- R files used for this course are available upon request.

Chapter 12:

Parameter Tests for Two Populations

12.1 Introduction

The context.

- Goal: Show methods to compare two populations.
- Each population is represented by a probability distribution.
- Then, we want to test a hypothesis about certain properties of the distributions.

12.1 Introduction

Example 1: Customer expenditure at a supermarket.

- Is there a difference between average expenditure of female and male customers of a supermarket?
- Here, there are two populations:

  population 1: female customers,

  population 2: male customers.
- Define:

$$\begin{aligned} X &= \text{expenditure of a female customer,} \\ Y &= \text{expenditure of a male customer.} \end{aligned}$$

- We should test:

$$H_0 : \mathrm{E}(X) = \mathrm{E}(Y) \quad \text{against} \quad H_1 : \mathrm{E}(X) \neq \mathrm{E}(Y)$$

12.1 Introduction

Example 2: Performance evaluation.

- Employees are trained in a certain task. — Is the training program effective?
- There are two populations:

  population 1: employees before training,

  population 2: employees after training.
- Define:

$$
\begin{aligned}
X &= \text{time required to do the task before training,} \\
Y &= \text{time required to do the task after training.}
\end{aligned}
$$

- To show that training has a positive effect, test:

$$H_0 : \mathrm{E}(X) \leq \mathrm{E}(Y) \quad \text{against} \quad H_1 : \mathrm{E}(X) > \mathrm{E}(Y)$$

12.1 Introduction

Sampling.

- We need data from each population to test a hypothesis concerning both populations.
- There are two concepts for drawing a sample:
  - independent samples:
    Draw a sample from each population independently.
    (Sample sizes need not be equal.)
  - matched pairs:
    Choose sample members in pairs, one from each population.
    (This will reduce the variation in comparisons.)
- Which is possible in the previous examples?

12.1 Introduction

The general situation.

- X, Y: our variables of interest
- The distribution of X (Y) depends on a parameter θ_X (θ_Y).
- The parameters θ_X, θ_Y are unknown; we want to test a hypothesis concerning θ_X and θ_Y.
- Independent samples: $X_1, \ldots, X_{n_X}, \; Y_1, \ldots, Y_{n_Y}$
- Matched pairs: $(X_1, Y_1), \ldots, (X_n, Y_n)$

12.1 Introduction

Outlook on Chapter 12.

- 12.2 Comparing Population Means I

  Normal distribution, independent samples; variances known / unknown
- 12.3 Comparing Population Means II

  Normal distribution, matched pairs; variances unknown
- 12.4 Comparing Population Means III

  No distributional assumptions, but the samples need to be large.
- 12.5 Comparing Two Shares

  Shares or unknown success probabilities.
- 12.6 Comparing Population Variances

  Normal distribution, independent samples

12.2 Comparing Population Means I

Independent Samples.

- Here we assume: $X \sim \mathsf{N}(\mu_X, \sigma_X^2)$, $Y \sim \mathsf{N}(\mu_Y, \sigma_Y^2)$.

- To test a null hypothesis which compares μ_X and μ_Y, the difference $\bar{X} - \bar{Y}$ will be crucial.

- What is the distribution of $\bar{X} - \bar{Y}$?

$$\bar{X} \sim \mathsf{N}(\mu_X, \sigma_X^2/n_X), \quad \bar{Y} \sim \mathsf{N}(\mu_Y, \sigma_Y^2/n_Y)$$

- Since $X_1, \ldots, X_{n_X}$ and $Y_1, \ldots, Y_{n_Y}$ are independent,

$$\bar{X} - \bar{Y} \sim \mathsf{N}\left(\mu_X - \mu_Y, \frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}\right)$$

12.2 Comparing Population Means I

Independent Samples.

- If $\mu_X = \mu_Y$, standardization leads to:

$$\frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}}} \sim \mathsf{N}(0, 1).$$

- The standardization of $\bar{X} - \bar{Y}$ is useful only if the parameters σ_X^2 and σ_Y^2 are known.

12.2 Comparing Population Means I

Are σ_X^2 and σ_Y^2 known?

- We substitute estimates if σ_X^2 and σ_Y^2 are unknown.

- Three cases:

  - σ_X^2, σ_Y^2 known
  - σ_X^2, σ_Y^2 unknown, but assumed equal
  - σ_X^2, σ_Y^2 unknown, not even assumed equal

- We shall now see these cases in some more detail.

12.2 Comparing Population Means I

Testing for the means when the variances are known.

- $X \sim \mathrm{N}(\mu_X, \sigma_X^2)$, $Y \sim \mathrm{N}(\mu_Y, \sigma_Y^2)$; σ_X^2 and σ_Y^2 known
- Test problem: $H_0 : \mu_X = \mu_Y$, $H_1 : \mu_X \neq \mu_Y$
- Test statistic:

$$T = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{\sigma_X^2}{n_X} + \frac{\sigma_Y^2}{n_Y}}}$$

  If H_0 is true, $T \sim \mathrm{N}(0, 1)$.

- Critical for H_0: too small and too large values of T.
- For the one-sided problem $H_0 : \mu_X \leq \mu_Y$, $H_1 : \mu_X > \mu_Y$, too large values of T are critical.

12.2 Comparing Population Means I

Testing for the means when the variances are unknown.

- $X \sim \mathrm{N}(\mu_X, \sigma_X^2)$, $Y \sim \mathrm{N}(\mu_Y, \sigma_Y^2)$; σ_X^2, σ_Y^2 unknown but assumed equal
- Test problem: $H_0 : \mu_X = \mu_Y$, $H_1 : \mu_X \neq \mu_Y$
- Test statistic (with the "pooled variance" s_p^2):

$$T = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{s_p^2}{n_X} + \frac{s_p^2}{n_Y}}} \quad \text{with} \quad s_p^2 = \frac{(n_X - 1)s_X^2 + (n_Y - 1)s_Y^2}{n_X + n_Y - 2}$$

  If H_0 is true, $T \sim t_{n_X + n_Y - 2}$.

- Critical for H_0: too small and too large values of T.
- For the one-sided problem $H_0 : \mu_X \leq \mu_Y$, $H_1 : \mu_X > \mu_Y$, too large values of T are critical.

12.2 Comparing Population Means I

Example: Comparing two varieties of oranges.

- Two varieties A and B of oranges should be compared with respect to their average juice content.

- To be tested: $H_0 : \mu_A = \mu_B$ against $H_1 : \mu_A \neq \mu_B$

- Data:

brand	sample size	average juice content	standard deviation
A	10	79.3 ml	12.3 ml
B	15	68.3 ml	14.1 ml

- This test problem will now be carried out under different assumptions.

12.2 Comparing Population Means I

Example: Comparing two varieties of oranges.

- Data:

brand	sample size	average juice content	standard deviation
A	10	79.3 ml	12.3 ml
B	15	68.3 ml	14.1 ml

- Now suppose we know: $\sigma_X^2 = \sigma_Y^2 = 13^2$.
- The test statistic is then:

$$T = \frac{79.3 - 68.3}{\sqrt{\frac{13^2}{10} + \frac{13^2}{15}}} = 2.07$$

- With $\alpha = 5\%$, the null hypothesis will be rejected:
  We found a significant difference between the two varieties.

12.2 Comparing Population Means I

Example: Comparing two varieties of oranges.

- Data:

brand	sample size	average juice content	standard deviation
A	10	79.3 ml	12.3 ml
B	15	68.3 ml	14.1 ml

- Now suppose all we know is: $\sigma_X^2 = \sigma_Y^2$.
- Then $s_p^2 = \frac{9 \cdot 12.3^2 + 14 \cdot 14.1^2}{10+15-2} = 180.2$ and the test statistic is:

$$T = \frac{79.3 - 68.3}{\sqrt{\frac{180.2}{10} + \frac{180.2}{15}}} = 2.01$$

- Since $t_{0.975;23} = 2.07$, the null hypothesis will not be rejected: No significant difference between the two varieties found.

12.2 Comparing Population Means I

Example: Comparing two varieties of oranges.

Some final remarks.

- What we find in data depends on what we already know.
- To decide if $H_0 : \mu_X = \mu_Y$ should be rejected, the difference $\bar{X} - \bar{Y}$ is not enough:
  Variances have to be taken into account.
- In practice, much care has to be taken to justify an assumption.
  (This is an ethical problem.)

12.2 Comparing Population Means I

Testing for the means when the variances are unknown.

- $X \sim \mathrm{N}(\mu_X, \sigma_X^2)$, $Y \sim \mathrm{N}(\mu_Y, \sigma_Y^2)$; σ_X^2, σ_Y^2 unknown, not even assumed equal
- Test problem: $H_0 : \mu_X = \mu_Y$, $H_1 : \mu_X \neq \mu_Y$
- Test statistic:

$$T = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{s_X^2}{n_X} + \frac{s_Y^2}{n_Y}}}$$

- What is the distribution of T in the case where $\mu_X = \mu_Y$?

12.2 Comparing Population Means I

Testing for the means when the variances are unknown.

- This problem is known as the Behrens-Fisher problem. Only approximate solutions are available.

- Method of Welch: If $\mu_X = \mu_Y$, then approximately $T \sim t_k$ where k is the largest integer not larger than

$$k = \frac{\left(\frac{s_X^2}{n_X} + \frac{s_Y^2}{n_Y}\right)^2}{\frac{1}{n_X - 1}\left(\frac{s_X^2}{n_X}\right)^2 + \frac{1}{n_Y - 1}\left(\frac{s_Y^2}{n_Y}\right)^2}$$

- If $\mu_X = \mu_Y$ and $n_X, n_Y > 30$, $T \sim \mathrm{N}(0, 1)$ approximately.

12.3 Comparing Population Means II

Matched pairs.

- Again, we assume: $X \sim \mathsf{N}(\mu_X, \sigma_X^2)$, $Y \sim \mathsf{N}(\mu_Y, \sigma_Y^2)$.

- Using matched pairs, we can test a null hypothesis about μ_X and μ_Y in terms of $D = X - Y$.

- Based on a sample $(X_1, Y_1), \ldots, (X_n, Y_n)$, define:

$$D_i = X_i - Y_i, \quad \bar{D} = \frac{1}{n} \sum_{i=1}^{n} D_i$$

- Then, with σ_D^2 being the variance of $D = X - Y$,

$$\bar{D} \sim \mathsf{N}(\mu_X - \mu_Y, \sigma_D^2 / n).$$

12.3 Comparing Population Means II

The advantage of matched pairs.

- Taking matched pairs may lead to a reduction in the variance of the difference.
- If X and Y are independent: $\sigma_D^2 = \sigma_X^2 + \sigma_Y^2$, and there is no advantage.
- The advantage is due to the dependence (correlation) of X and Y; for example:
  Someone already very skilled before training will again be very skilled after training.

12.3 Comparing Population Means II

Testing for the means, matched pairs.

- $X \sim \mathsf{N}(\mu_X, \sigma_X^2)$, $Y \sim \mathsf{N}(\mu_Y, \sigma_Y^2)$; σ_X^2 and σ_Y^2 unknown
- Test problem: $H_0 : \mu_X = \mu_Y$, $H_1 : \mu_X \neq \mu_Y$
- Test statistic:

$$T = \frac{\bar{D}}{\sqrt{\frac{s_D^2}{n}}}, \quad \text{where} \quad s_D^2 = \frac{1}{n-1}\sum_{i=1}^{n}(D_i - \bar{D})^2$$

  If H_0 is true, $T \sim t_{n-1}$.
- Critical for H_0: too small and too large values of T.
- For the one-sided problem $H_0 : \mu_X \leq \mu_Y$, $H_1 : \mu_X > \mu_Y$, too large values of T are critical.

12.3 Comparing Population Means II

Example: Improving forecasting capabilities.

- Five students were asked to forecast the next day's closing value of İMKB 100 twice:
  - before a training phase ("time 1")
  - after a training phase ("time 2")
- Their relative forecast errors in percent (taken positive) were:

student #	1	2	3	4	5
X = relative error, time 1	2.1	1.8	3.1	2.1	2.5
Y = relative error, time 2	0.3	0.7	0.4	0.1	0.2
$D = X - Y$	1.8	1.1	2.7	2.0	2.3

- Is there evidence that the students could improve their forecasting capabilities during the training phase?

12.3 Comparing Population Means II

Example: Improving forecasting capabilities.

- To show the positive training effect, test

$$H_0 : \mu_X \leq \mu_Y \quad \text{against} \quad H_1 : \mu_X > \mu_Y$$

- For the data set above, $\bar{D} = 1.98$, $s_D^2 = 0.357$, $n = 5$
- The test statistic is

$$T = \frac{1.98}{\sqrt{\frac{0.357}{5}}} = 7.41$$

- Since $t_{0.95;4} = 2.132$, H_0 is rejected.
- Conclusions?

12.4 Comparing Population Means III

Testing for population means, large independent samples.

- Test problem: $H_0 : \mu_X = \mu_Y$, $H_1 : \mu_X \neq \mu_Y$
  ($\mu_X = \mathrm{E}(X)$, $\mu_Y = \mathrm{E}(Y)$)
- Test statistic:
$$T = \frac{\bar{X} - \bar{Y}}{\sqrt{\frac{s_X^2}{n_X} + \frac{s_Y^2}{n_Y}}}$$
  If H_0 is true, $T \sim \mathrm{N}(0, 1)$ approximately, if n_X, n_Y are large.
- Critical for H_0: too small and too large values of T.
- For the one-sided test problem $H_0 : \mu_X \leq \mu_Y$,
  $H_1 : \mu_X > \mu_Y$, too large values of T are critical.

12.4 Comparing Population Means III

Example: Customer expenditure at a supermarket.

- We'll pick up on an example we've seen earlier.
- Is there a difference between average expenditure of female and male customers of a supermarket?
- Two populations: female / male customers
- With

$$\begin{aligned} X &= \text{expenditure of a female customer,} \\ Y &= \text{expenditure of a male customer,} \end{aligned}$$

we would like to test:

$$H_0 : \mathrm{E}(X) = \mathrm{E}(Y) \quad \text{against} \quad H_1 : \mathrm{E}(X) \neq \mathrm{E}(Y)$$

12.4 Comparing Population Means III

Example: Customer expenditure at a supermarket.

- There are two independent samples:

$$\begin{array}{lll} X_1, \ldots, X_{366} & \bar{x} = 15.19 & s_X^2 = 162.68 \\ Y_1, \ldots, Y_{142} & \bar{y} = 16.06 & s_Y^2 = 178.68 \end{array}$$

- The test statistic is:

$$T = \frac{15.19 - 16.06}{\sqrt{\frac{162.68}{366} + \frac{178.68}{142}}} = -0.67$$

- The (approximate) p-value is

$$P_{H_0 \text{ true}}(T < -0.67 \text{ or } T > +0.67) = 0.503.$$

12.4 Comparing Population Means III

Example: Customer expenditure at a supermarket.

- This means: $H_0 : \mathrm{E}(X) = \mathrm{E}(Y)$ is not rejected against $H_1 : \mathrm{E}(X) \neq \mathrm{E}(Y)$.
- So, why have we observed a difference between average expenditure of female and male customers?
- What could our result imply for the management of the supermarket?
- What else might the management be interested in?

12.5 Comparing Two Shares

Approximate test if two success probabilities are equal.

- Test problem: $H_0 : p_X = p_Y$, $H_1 : p_X \neq p_Y$
- Test statistic:

$$T = \frac{\hat{p}_X - \hat{p}_Y}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n_X} + \frac{\hat{p}(1-\hat{p})}{n_Y}}} \quad \text{with} \quad \hat{p} = \frac{n_X \hat{p}_X + n_Y \hat{p}_Y}{n_X + n_Y}$$

  If H_0 is true, $T \sim \mathrm{N}(0, 1)$ approximately, if n_X, n_Y are large.

- Critical for H_0: too small and too large values of T.
- For the one-sided test problem $H_0 : p_X \leq p_Y$, $H_1 : p_X > p_Y$, too large values of T are critical.

12.5 Comparing Two Shares

Example: Brand awareness ratings.

- A cosmetics company commissions a marketing campaign to build brand awareness in the market.
- Before the campaign, 238 in a survey of 750 said they know the brand.
- After the campaign, 361 in a new survey of 1050 said they know the brand.
- Is there evidence that the marketing campaign was successful?

12.5 Comparing Two Shares

Example: Brand awareness ratings.

- Define:

$$p_X = \text{brand awareness rating before the campaign,}$$
$$p_Y = \text{brand awareness rating after the campaign.}$$

- To show that the campaign was successful, we have to test

$$H_0 : p_X \geq p_Y \quad \text{against} \quad H_1 : p_X < p_Y .$$

- For this test problem too small values of the test statistic T are critical.

12.5 Comparing Two Shares

Example: Brand awareness ratings.

- Here, we have:

$$\hat{p}_X = \frac{238}{750}, \quad \hat{p}_Y = \frac{361}{1050}, \quad \hat{p} = \frac{238+361}{750+1050} = \frac{599}{1800}$$

- Test statistic:

$$T = \frac{\frac{238}{750} - \frac{361}{1050}}{\sqrt{\frac{\frac{599}{1800}(1-\frac{599}{1800})}{750} + \frac{\frac{599}{1800}(1-\frac{599}{1800})}{1050}}} = -1.18$$

- With $\alpha = 5\%$, the null hypothesis will not be rejected: No significant difference between the two shares found.

12.6 Comparing Population Variances

Testing for the variances, independent samples.

- $X \sim \mathsf{N}(\mu_X, \sigma_X^2)$, $Y \sim \mathsf{N}(\mu_Y, \sigma_Y^2)$
- Test problem: $H_0 : \sigma_X^2 = \sigma_Y^2$, $H_1 : \sigma_X^2 \neq \sigma_Y^2$
- Test statistic:

$$T = \frac{s_X^2}{s_Y^2}$$

  If H_0 is true, $T \sim F_{n_X - 1, n_Y - 1}$.

- Critical for H_0: too small and too large values of T.
- For the one-sided problem $H_0 : \sigma_X^2 \leq \sigma_Y^2$, $H_1 : \sigma_X^2 > \sigma_Y^2$, too large values of T are critical.

12.6 Comparing Population Variances

Example: The variances of response times.

- A computerized inquiry system provides information on the availability of air tickets.
- As long as the system is running smoothly, response times are known to be normally distributed.
- The variance of response times is an important performance parameter.
- Is there evidence that the variances in the night and in the daytime differ?

12.6 Comparing Population Variances

Example: The variances of response times.

- Define:

$$X = \text{response time in the night,}$$
$$Y = \text{response time in the daytime.}$$

- One day, two samples were taken:
  - A sample of X of size 384 had standard deviation $s_X = 0.1085$.
  - A sample of Y of size 352 had standard deviation $s_Y = 0.1115$.

- Is this evidence enough that the variances of X and Y differ?

12.6 Comparing Population Variances

Example: The variances of response times.

- Test problem: $H_0 : \sigma_X^2 = \sigma_Y^2$, $H_1 : \sigma_X^2 \neq \sigma_Y^2$
- Test statistic:

$$T = \frac{s_X^2}{s_Y^2} = \frac{0.1085^2}{0.1115^2} = 0.9469$$

- 2.5% and 97.5% quantiles of the F distribution with 383 numerator and 351 denominator degrees of freedom:

$$F_{383;351;0.025} = 0.8149; \quad F_{383;351;0.975} = 1.2283.$$

- Critical region: $[0, 0.8149) \cup (1.2283, \infty)$.
- H_0 is not rejected.