Chapter 10: Inferences based on two samples

MATH 450

November 28^{th}, 2017

Overview

Week 1	Chapter 1: Descriptive statistics
Week 2	Chapter 6: Statistics and Sampling Distributions
Week 4	Chapter 7: Point Estimation
Week 7	Chapter 8: Confidence Intervals
Week 10	Chapter 9: Tests of Hypotheses
Week 12	**Chapter 10: Two-sample inference**

Overview

Example

Let μ_1 and μ_2 denote true average decrease in cholesterol for two drugs. From two independent samples $X_1, X_2, \ldots, X_m$ and $Y_1, Y_2, \ldots, Y_n$, we want to test:

$$H_0 : \mu_1 = \mu_2$$
$$H_a : \mu_1 \neq \mu_2$$

Settings

- Independent samples
  1. $X_1, X_2, \ldots, X_m$ is a random sample from a population with mean μ_1 and variance σ_1^2.
  2. $Y_1, Y_2, \ldots, Y_n$ is a random sample from a population with mean μ_2 and variance σ_2^2.
  3. The X and Y samples are independent of each other.
- Today: paired samples
  1. There is only one set of n individuals or experimental objects
  2. Two observations are made on each individual or object

Independent samples

1. $X_1, X_2, \ldots, X_m$ is a random sample from a population with mean μ_1 and variance σ_1^2.
2. $Y_1, Y_2, \ldots, Y_n$ is a random sample from a population with mean μ_2 and variance σ_2^2.
3. The X and Y samples are independent of each other.

Testing the difference between two population means

- Setting: independent normal random samples $X_1, X_2, \ldots, X_m$ and $Y_1, Y_2, \ldots, Y_n$ with known values of σ_1 and σ_2. Constant Δ_0.
- Null hypothesis:

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

- Alternative hypothesis:
  (a) $H_a : \mu_1 - \mu_2 > \Delta_0$
  (b) $H_a : \mu_1 - \mu_2 < \Delta_0$
  (c) $H_a : \mu_1 - \mu_2 \neq \Delta_0$
- When $\Delta = 0$, the test (c) becomes

$$H_0 : \mu_1 = \mu_2$$
$$H_a : \mu_1 \neq \mu_2$$

Proposition

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

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

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

2-sample t test: degree of freedom

THEOREM When the population distributions are both normal, the standardized variable

$$T = \frac{\bar{X} - \bar{Y} - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{m} + \frac{S_2^2}{n}}} \tag{10.2}$$

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

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

where

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

(round ν down to the nearest integer).

2-sample t procedures

The **two-sample t test** for testing H_0: $\mu_1 - \mu_2 = \Delta_0$ is as follows:

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

Alternative Hypothesis	**Rejection Region for Approximate Level α Test**
H_a: $\mu_1 - \mu_2 > \Delta_0$	$t \geq t_{\alpha,\nu}$ (upper-tailed test)
H_a: $\mu_1 - \mu_2 < \Delta_0$	$t \leq -t_{\alpha,\nu}$ (lower-tailed test)
H_a: $\mu_1 - \mu_2 \neq \Delta_0$	either $t \geq t_{\alpha/2,\nu}$ or $t \leq -t_{\alpha/2,\nu}$ (two-tailed test)

A P-value can be computed as described in Section 9.4 for the one-sample t test.

Paired samples

1. There is only one set of n individuals or experimental objects
2. Two observations are made on each individual or object

Settings

Assumption

1. *The data consists of n independently selected pairs* $(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)$ *with* $E(X_i) = \mu_1$ *and* $E(Y_i) = \mu_2$.
2. *Let*

   $$D_1 = X_1 - Y_1, \qquad D_2 = X_2 - Y_2, \ldots, \qquad D_n = X_n - Y_n,$$

   so the D_i*'s are the differences within pairs.*
3. *The* D_i*'s are assumed to be normally distributed with mean value* μ_D *and variance* σ_D^2.

From the last lecture (independent samples)

Problem

Assume that

- $X_1, X_2, \ldots, X_m$ *is a random sample from a population with mean* μ_1 *and variance* σ_1^2.
- $Y_1, Y_2, \ldots, Y_n$ *is a random sample from a population with mean* μ_2 *and variance* σ_2^2.
- *The* X *and* Y *samples are independent of each other.*

Compute (in terms of $\mu_1, \mu_2, \sigma_1, \sigma_2, m, n$*)*

(a) $E[\bar{X} - \bar{Y}]$

(b) $Var[\bar{X} - \bar{Y}]$ *and* $\sigma_{\bar{X}-\bar{Y}}$

Notes

- In the independent case, we construct the statistics by looking at the distribution of

$$\bar{X} - \bar{Y}$$

  which have

$$E[\bar{X}-\bar{Y}] = \mu_1-\mu_2, \qquad Var[\bar{X}-\bar{Y}] = Var(\bar{X})+Var(\bar{Y}) = \frac{\sigma_1^2}{m}+\frac{\sigma_2^2}{n}$$

- With paired data, the X and Y observations within each pair are not independent, so $\bar{X}$ and $\bar{Y}$ are not independent of each other $\rightarrow$ the computation of the variance is in valid $\rightarrow$ could not use the old formulas

The paired t-test

- Because different pairs are independent, the D_i's are independent of each other
- We also have

$$E[D] = E[X - Y] = E[X] - E[Y] = \mu_1 - \mu_2 = \mu_D$$

- Testing about $\mu_1 - \mu_2$ is just the same as testing about μ_D
- Idea: to test hypotheses about $\mu_1 - \mu_2$ when data is paired:
  1. form the differences $D_1, D_2, \ldots, D_n$
  2. carry out a one-sample t-test (based on $n - 1$ df) on the differences.

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: $t = \dfrac{\overline{d} - \Delta_0}{s_D/\sqrt{n}}$ (where $\overline{d}$ 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: $\mu_D > \Delta_0$	$t \geq t_{\alpha,n-1}$
H_a: $\mu_D < \Delta_0$	$t \leq -t_{\alpha,n-1}$
H_a: $\mu_D \neq \Delta_0$	either $t \geq t_{\alpha/2,n-1}$ or $t \leq -t_{\alpha/2,n-1}$

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

Confidence intervals

- A t confidence interval for for $\mu_D = \mu_1 - \mu_2$ can be constructed based on the fact that

$$T = \frac{\bar{D} - \mu_D}{S_D/\sqrt{n}}$$

follows the t distribution with degree of freedom $n-1$.

- The paired t CI for μ_D is

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

- A one-sided confidence bound results from retaining the relevant sign and replacing $t_{\alpha/2,n-1}$ by $t_{\alpha,n-1}$.

Example

Example

Consider two scenarios:

A. Insulin rate is measured on 30 patients before and after a medical treatment.

B. Insulin rate is measured on 30 patients receiving a placebo and 30 other patients receiving a medical treatment.

What type of test should be used in each cased: paired or unpaired?

Example

Example

Suppose we have a new synthetic material for making soles for shoes. We'd like to compare the new material with leather – using some energetic kids who are willing to wear test shoes and return them after a time for our study. Consider two scenarios:

A. Giving 50 kids synthetic sole shoes and 50 kids leather shoes and then collect them back, comparing the average wear in each group

B. Give each of a random sample of 50 kids one shoe made by the new synthetic materials and one shoe made with leather

What type of test should be used in each cased: paired or unpaired?

Example

Consider an experiment in which each of 13 workers was provided with both a conventional shovel and a shovel whose blade was perforated with small holes. The following data on stable energy expenditure is provided:

Worker:	1	2	3	4	5	6	7
Conventional:	.0011	.0014	.0018	.0022	.0010	.0016	.0028
Perforated:	.0011	.0010	.0019	.0013	.0011	.0017	.0024

Worker:	8	9	10	11	12	13
Conventional:	.0020	.0015	.0014	.0023	.0017	.0020
Perforated:	.0020	.0013	.0013	.0017	.0015	.0013

Calculate a confidence interval at the 95% confidence level for the true average difference between energy expenditure for the conventional shovel and the perforated shovel (assuming that the differences follow normal distribution).

Example

Consider an experiment in which each of 13 workers was provided with both a conventional shovel and a shovel whose blade was perforated with small holes. The following data on stable energy expenditure is provided:

Worker:	1	2	3	4	5	6	7
Conventional:	.0011	.0014	.0018	.0022	.0010	.0016	.0028
Perforated:	.0011	.0010	.0019	.0013	.0011	.0017	.0024

Worker:	8	9	10	11	12	13
Conventional:	.0020	.0015	.0014	.0023	.0017	.0020
Perforated:	.0020	.0013	.0013	.0017	.0015	.0013

Carry out a test of hypotheses at significance level .05 to see if true average energy expenditure using the conventional shovel exceeds that using the perforated shovel; include a P-value in your analysis.

t-table

t \ ν	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1.6	.178	.125	.104	.092	.085	.080	.077	.074	.072	.070	.069	.068	.067	.065	.065	.065	.064	.064
1.7	.169	.116	.094	.082	.075	.070	.065	.064	.062	.060	.059	.057	.056	.055	.055	.054	.054	.053
1.8	.161	.107	.085	.073	.066	.061	.057	.055	.053	.051	.050	.049	.048	.046	.046	.045	.045	.044
1.9	.154	.099	.077	.065	.058	.053	.050	.047	.045	.043	.042	.041	.040	.038	.038	.038	.037	.037
2.0	.148	.092	.070	.058	.051	.046	.043	.040	.038	.037	.035	.034	.033	.032	.032	.031	.031	.030
2.1	.141	.085	.063	.052	.045	.040	.037	.034	.033	.031	.030	.029	.028	.027	.027	.026	.025	.025
2.2	.136	.079	.058	.046	.040	.035	.032	.029	.028	.026	.025	.024	.023	.022	.022	.021	.021	.021
2.3	.131	.074	.052	.041	.035	.031	.027	.025	.023	.022	.021	.020	.019	.018	.018	.018	.017	.017
2.4	.126	.069	.048	.037	.031	.027	.024	.022	.020	.019	.018	.017	.016	.015	.015	.014	.014	.014
2.5	.121	.065	.044	.033	.027	.023	.020	.018	.017	.016	.015	.014	.013	.012	.012	.012	.011	.011
2.6	.117	.061	.040	.030	.024	.020	.018	.016	.014	.013	.012	.012	.011	.010	.010	.010	.009	.009
2.7	.113	.057	.037	.027	.021	.018	.015	.014	.012	.011	.010	.010	.009	.008	.008	.008	.008	.007
2.8	.109	.054	.034	.024	.019	.016	.013	.012	.010	.009	.009	.008	.008	.007	.007	.006	.006	.006
2.9	.106	.051	.031	.022	.017	.014	.011	.010	.009	.008	.007	.007	.006	.005	.005	.005	.005	.005
3.0	.102	.048	.029	.020	.015	.012	.010	.009	.007	.007	.006	.006	.005	.004	.004	.004	.004	.004