Module 5. Bivariate analysis: qualitative — quantitative

Data Science & AI

Sabine De Vreese   Lieven Smits   Bert Van Vreckem

2023–2024

HO
GENT

Contents

Data visualization

Two-sample t-test

Paired samples

Effect size

HO
GENT

Learning goals

- Apply the *t*-test for two samples
- Calculate effect size using Cohen's *d*
- Visualization

HO
GENT

Bivariate analysis: overview

Independent	**Dependent**	**Test/Metric**
Qualitative	Qualitative	χ^2-test Cramér's V
Qualitative	Quantitative	two-sample t-test Cohen's d
Quantitative	Quantitative	— Regression, correlation

HO GENT

Example research questions

- Are male penguins larger than females?
- Do men get a higher salary than women?
- Does a new vaccine protect against a disease?
- Does “retrieval practice” improve learning outcomes (i.e. student grades)?
- …

In these examples, what is the independent/dependent variable?

HO
GENT

Data visualization

HO
GENT

Data visualization

- Chart types for quantitative data
- Grouped by qualitative variable

HO
GENT

Data visualization

- Chart types for quantitative data
- Grouped by qualitative variable

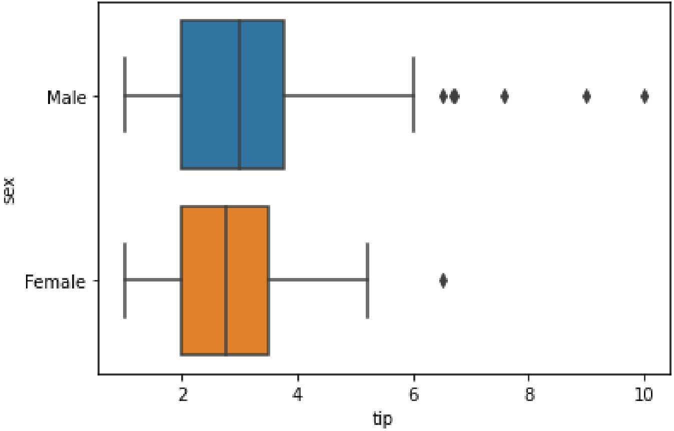
Suitable chart types:

- Grouped boxplot
- Grouped density plot
- Bar chart with error bars
- ...

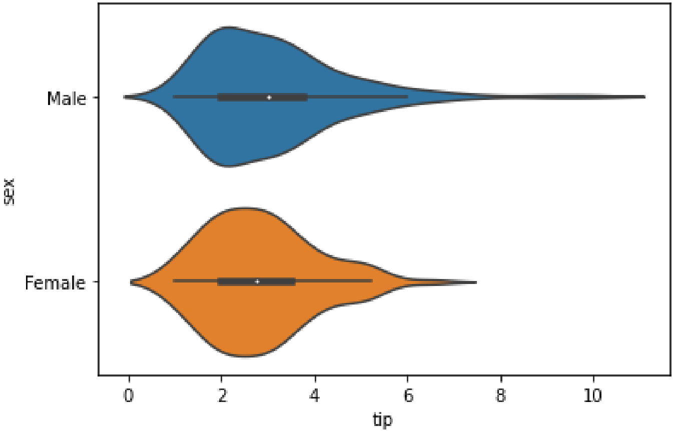
HO
GENT

Grouped boxplot

(source code: see `demo-5.01-2sample-t.ipynb`)

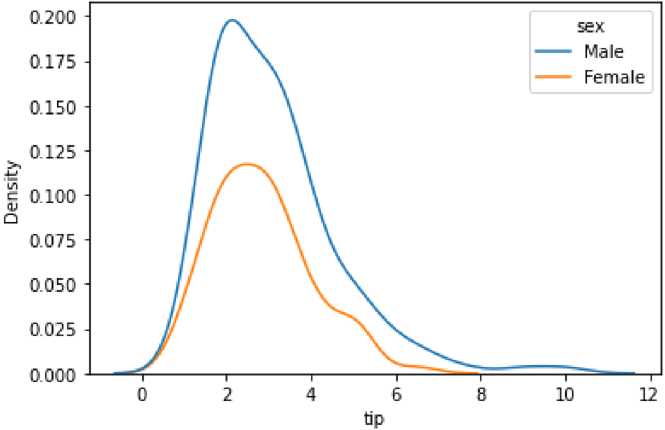


Grouped violin plot



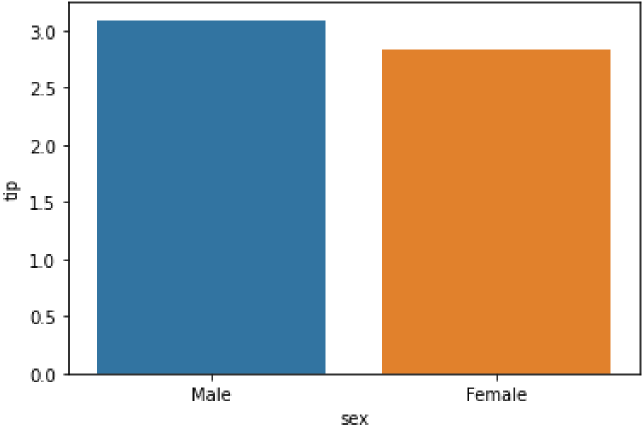
HO
GENT

Grouped density plot



HO
GENT

Beware! Bar chart of group means



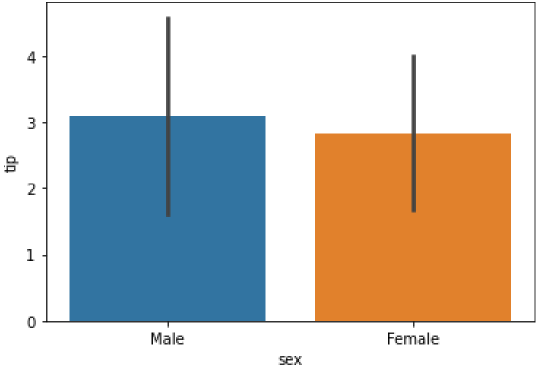
HO GENT

Beware! Bar chart of group means

HO
GENT

Bar chart with error bars

- Always add error bars!
- Only makes sense for normally distributed data
- Example: 1 standard deviation:



HO
GENT

Two-sample t-test

HO
GENT

Comparing two population means

Are the population means of two populations different?

We use two *samples* to perform an appropriate statistical test.

Correct test depends on:

- Independent samples
- Paired samples

HO
GENT

Independent samples

In a clinical study, the aim is to determine whether a new drug has a delayed (i.e. higher) reaction time as a side effect.

- Control group: 6 participants receive placebo
- Intervention group: 6 participants receive medicine

HO
GENT

Independent samples

In a clinical study, the aim is to determine whether a new drug has a delayed (i.e. higher) reaction time as a side effect.

- Control group: 6 participants receive placebo
- Intervention group: 6 participants receive medicine

Next, the reaction time (in ms) is measured:

- Control group: 91, 87, 99, 77, 88, 91 ($\overline{x}$ = 88.83)
- Intervention group: 101, 110, 103, 93, 99, 104 ($\overline{y}$ = 101.67)

Are there significant differences between the intervention and control group?

HO
GENT

Independent samples

Testing procedure

1. Hypotheses:
   - $H_0: \mu_1 - \mu_2 = 0$
   - $H_1: \mu_1 - \mu_2 < 0$
2. Significance level: $\alpha = 0.05$
3. Test statistic is based on:
   - $\overline{x} - \overline{y} = -12.833$
   - $\overline{x}$ = estimation for μ_1 (control group)
   - $\overline{y}$ = estimation for μ_2 (intervention group)
   - Takes the variances of the samples into account. For completeness the test statistic is

$$t = \frac{\overline{x} - \overline{y}}{\sqrt{s_x^2/n_x + s_y^2/n_y}}.$$

4. Calculate p

HO
GENT

Independent samples

Calculation in Python

The calculation of the test statistic and the associated *p*-value is done by `stats.ttest_ind()`

```
control = np.array([91, 87, 99, 77, 88, 91])
treatment = np.array([101, 110, 103, 93, 99, 104])
stats.ttest_ind(a=control, b=treatment,
                alternative='less', equal_var=False)
```

Result:

```
Ttest_indResult(statistic=-3.445612673536487,
  pvalue=0.003391230079206901)
```

HO
GENT

Independent samples

Testing procedure (continued)

5. Draw conclusion based on p-value.
   $p \approx 0.00339 < \alpha = 0.05$. We reject the null hypothesis. In this sample, there is reason to assume that the drug does indeed increase reaction time.

HO
GENT

Paired samples

HO
GENT

Paired samples

A study examined whether cars that run on petrol with additives have a lower consumption, i.e. a higher miles per gallon.
For 10 cars, the consumption was measured (expressed in miles per gallon) for both fuel types:

Car	1	2	3	4	5	6	7	8	9	10
Regular petrol	16	20	21	22	23	22	27	25	27	28
With additives	19	22	24	24	25	25	25	26	28	32

HO
GENT

Paired samples

Testing procedure

1. Hypotheses:
   - H_0: $\mu_{X-Y} = 0$
   - H_1: $\mu_{X-Y} < 0$
2. Significance level: $\alpha = 0.05$
3. Test statistic:
   - Based on the mean of the difference $\overline{d} = \overline{x - y}$

Regular petrol	16	20	21	22	23	22	27	25	27	28
With additives	19	22	24	24	25	25	25	26	28	32
Difference	-3	-2	-3	-2	-2	-3	2	-1	-1	-4

   - For completeness, the test statistic is

$$t = \frac{\overline{d}}{s_d / \sqrt{n}}$$

4. Calculate p

HO GENT

Paired samples

Calculation in Python

```
regular = np.array([16, 20, 21, 22, 23, 22, 27, 25, 27, 28])
additives = np.array([19, 22, 24, 24, 25, 25, 26, 26, 28, 32])
stats.ttest_rel(regular, additives, alternative='less')
```

Result:

```
Ttest_relResult(statistic=-4.47213595499958,
    pvalue=0.0007749429585091)
```

HO
GENT

Paired samples

Testing procedure

5. Draw conclusion based on p-value.
   $p \approx 0.0007749 < \alpha = 0.05$. We reject the null hypothesis. In this sample, there is reason to assume that the fuel with additives leads to lower fuel consumption, i.e. higher miles per gallon.

HO
GENT

Effect size

HO
GENT

Effect size

Effect size

The effect size is a metric which expresses how great the difference between two groups is

- Control group vs. intervention group
- Can be used in addition to hypothesis test
- Often used in educational sciences
- There are several definitions, here: *Cohen's d*

HO GENT

Cohen's *d*

$$d = \frac{\overline{x_1} - \overline{x_2}}{s}$$

where $\overline{x_1}$ and $\overline{x_2}$ represent the sample means
and s the pooled standard deviation, i.e. standard deviation of both groups combined:

$$s = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$$

with n_1, n_2 the sample sizes, and s_1, s_2 the standard deviations of both groups

HO
GENT

Interpretation Cohen's *d*

$\lvert d \rvert$	**Effect Size**
0.01	Very small
0.2	Small
0.5	Average
0.8	Large
1.2	Very large
2.0	Huge

In educational sciences (John Hattie):

- 0,4 = tipping point for desired effects
- effect size ***d* = 1**: process material that would normally take 1 year in 6 months!

E.g. `https://visible-learning.org/backup-hattie-ranking-256-effects-2017/`

HO GENT

Typical approach research in education

- Research question: Is X a good learning strategy, in other words, does this have a positive effect on final results?
- Control group uses “traditional” approach
- Intervention group uses X
- Followed by an evaluation moment
- Determine scores, calculate *d*

HO
GENT