Thomas Haslwanter

An Introduction to Statistics With Python

– Errata –

January 2, 2018

Springer

Background

On p. 78, **Eq. 5.4** should be

For discrete distributions, the integral over x is replaced by the sum over all possible values:

$$E[X] = \sum_i x_i P_i \,. \tag{0.1}$$

where x_i represents all possible values that the measured variable can have.

Distributions of One Variable

On p. 108, the wrong figure has been inserted. **Fig. 6.11** should be

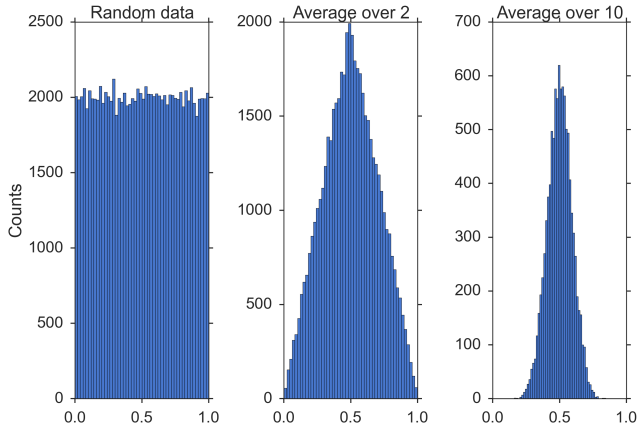

Fig. 0.1 Demonstration of the Central Limit Theorem for a uniform distribution: Left) Histogram of uniformly distributed random data between 0 and 1. Center) Histogram of average over two data points.) Right) Histogram of average over 10 data points.

Tests of Means of Numerical Data

- On p. 155, the **Summary: Selecting the Right Test for Comparing Groups** should read

No. of Groups Compared	**Independent Samples**	**Paired Samples**
Groups of Nominal Data		
2 or more	Fisher's exact test or Chi-Square test	McNemar's test
Groups of Ordinal Data		
2	Mann-Whitney U test	Wilcoxon signed rank test
3 or more	Kruskal-Wallis test	Friedman test
Groups of Continuous Data		
1	one-sample t-test or Wilcoxon signed rank sum test	—
2	Student's t-test or Mann-Whitney test	Paired t-test or Wilcoxon signed-rank sum test
3 or more	ANOVA or Kruskal-Wallis test	Repeated Measures ANOVA or Friedman test

Table 0.1 Typical tests for statistical problems, for nominal and ordinal data. Note that the tests for comparing one group to a fixed value are the same as comparing two groups with paired samples.

Hypothetical Examples

1 group, nominal Average calory intake. E.g. "Do our children eat more than they should?"

1 group, ordinal Sequence of giant-planets. E.g. "In our solar system, are giant planets further out than average in the sequence of planets?"

2 groups, nominal male/female, blond-hair/black-hair. E.g. "Are females more blond than males?"

2 groups, nominal, paired 2 labs, analysis of blood samples. E.g. "Does the blood analysis from Lab1 indicate more infections than the analysis from Lab2?"

2 groups, ordinal Jamaican/American, ranking 100m sprint. E.g. "Are Jamaican sprinters more successful than American sprinters?"

2 groups, ordinal, paired sprinters, before/after diet. E.g. "Does a chocolate diet make sprinters more successful?"

3 groups, ordinal single/married/divorces, ranking 100m sprint. E.g. "Does the marital status have an effect on the success of sprinters?"

3 groups, ordinal, paired sprinters, before/after diet. E.g. "Does a rice diet make Chinese sprinters more successful?"

2 groups, continuous male/female, IQ. E.g. "Are women more intelligent than men?"

2 groups, continuous, paired male/female, looking at diamonds. E.g. "Does looking at sports cars raise the male heart-beat more than the female?

3 groups, continuous Tyrolians, Viennese, Styrians; IQ. E.g. "Are Tyrolians smarter than people from other Austrian federal states?"

3 groups, continuous, paired Tyrolians, Viennese, Styrians; looking at mountains. E.g. "Does looking at mountains raise the heartbeat of Tyrolians more than those of other people?"

- On p. 157, The heading of the first exercise in 8.1 should be
  One sample t-test for the mean and **Wilcoxon signed rank sum test**

Tests on Categorical Data

- On p. 170, in Table 9.5, *Subject 9* should be listed only once.
- On p. 171, the code line `obs = [[a,b], [c, d]]` is wrong. However, the ISP-Quantlet listed there is correct.