Statistik II - Exercise session 3
12.11.2014 & 19.11.2014

Info

- Classroom: SPA1 203
- Time: Wednesdays, 16:15 - 17:45
- in English
- Assignments on webpage (lvb>staff>PB)

Contact: Petra Burdejova
petra.burdejova@hu-berlin.de
Office: SPA1 R400 (upon agreement)

Schedule:

Date	Week	Exercises
15.10.14	E1	4-11, 4-18, 4-19
22.10.14	E1	4-11, 4-18, 4-19
29.10.14	E2	5-6, 5-10, 5-11, 5-14 ,5-16
5.11.14	E2	5-6, 5-10, 5-11, 5-14 ,5-16
12.11.14	E3	5-16, 5-21, 5-23, 6-1
19.11.14	E3	5-16, 5-21, 5-23, 6-1
26.11.14	E4	6-3, 6-9, 6-13
03.12.14	E4	6-3, 6-9, 6-13
10.12.14	E5	TBA
17.12.14	E5	TBA
07.01.15	E6	TBA
14.01.15	E6	TBA
21.01.15	E7	TBA
28.01.15	–	–
04.02.15	E8	Review for exam
11.02.15	E8	Review for exam

Review

- week 5 & week 6
- Slides:
  - Distributions (04_Verteilungsmodelle) (1-50)
  - Theory of sampling (05_Stichprobentheorie) (1-65)

Distributions of random variables

See previous exercise session (SET2)
topics: discrete and continuous distributions, specific normal distribution ...

Theory of sampling

X - variable
x_i, $i = 1, \ldots, N$ elements of population (N - size of population)
x_j, $j = 1, \ldots, K$ values of variable with frequencies $f(x_j)$

mean $\mu = \frac{1}{N}\sum_{i=1}^{n} x_i = \frac{1}{N}\sum_{j=1}^{K} x_j f(x_j)$
variance $\sigma^2 = \frac{1}{N}\sum_{i=1}^{n}(x_i - \mu)^2 = \frac{1}{N}\sum_{j=1}^{K}(x_j - \mu)^2 f(x_j)$

Sample = finite subset of element from population
n - denote the size of sample

Sampling variables $X_1, \ldots, X_n$

Sampling function $U = U(X_1, \ldots, X_n)$

(U as function of r.variables is r.v. itself!)
Sample mean $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$
Sample variance

$$S^{*2} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \mu)^2$$
$$S'^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \overline{X})^2$$
$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$$

Because $\mathrm{E}[\bar{X}] = \mu$ and $\mathrm{E}[S^2] = \sigma^2$, they are unbiased estimates of μ and σ^2.

Proof of properties: see last page.

After drawing a sample we have concrete values $x_1, \ldots, x_n$.
$\bar{X}$ is r.v, however $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ is number!

Example:
X - Age of Berliner
N - Population
We are not able to ask everybody and compute μ - average age.
We can ask, make sample $X_1, \ldots, X_n$. Theoretically, it has specific properties, f.e. $\bar{X}$ is unbiased.
After asking, we have specific realizations $x_1, \ldots, x_n$ of $X_1, \ldots, X_n$.
Thus,if this sample is big enough, our estimate for μ will be ”good enough”.

Exercises

Exercise 5-16 - Steel pins

Machine produces steel pins. Unfortunately, diameter of pins fluctuates. R.v. "diameter of a pin" X_1 is normally distributed with $\mu = 6$mm and $\sigma = 0.4$mm.

a) What is the probability that diameter of pin is more than 2% away from $\mu = 6$mm?

b) What is the probability that diameter of pin is exactly 6mm?

c) Which value will not be exceeded with probability of 85%?

Another machine, which works independently of the first one, drilling holes into a work piece, in which the steel pins will be used. Also the diameter of holes is normally distributed random variable X_2 with $\mu = 6.05$mm and $\sigma = 0.3$mm.

d) What is the probability that diameter of hole is less than 6mm?

e) What is the distribution of $Y = X_2 - X_1$?

f) What is the probability that pin does not fit in the hole?

Exercise 5-21 - Archer

Archer A. Mor wants to impress his friends with his skill. He knows that he achieves on average 3 hits in 5 shots. To show it, he shoots 8 times on a target.

a) What is the distribution of r.v. X_i "Hit in the i-th shot" and r.v. Y "Total number of hits of all 8 shots"? Give reasons for your answer.

b) Give the expected number of hits in general and specifically for this example.

c) What is the probability that A. Mor achieves exactly three times?

Exercise 5-23 - Restaurant

In a restaurant, where only single contact-seeking guests can enter, a guest can make a choice from great menu.

A) On an ordinary Sunday, when restaurant is open from 19:00 till 24:00, usually 25 guest come - independently on each other. One can assume that the same number of guests is expected in each hour during the opening time.

   a) What is the probability that exactly one guest arrives in the first hour?

   b) How many minutes on average are between the arrivals of two guest on Sunday evening?

B) A guest has the opportunity to order again as often as he wants. The host assumes that each guest - independent of the other guests - orders at most once. He knows also that a guest does not reorder with 70% probability.

   a) What is the distribution of r.v. X "Number of extra orders on Sunday evening " precisely, if we assume usual number guests (as above)?

   b) If the ordinary number of guests comes, the probability that the capacity of the kitchen is exceeded by the reorders is 0.05%. How many guests have to reorder so that the capacity limit is reached?

Exercise 6-1 - Three people

From the population of 3 people $N = 3$ of age 20,22 and 24 was sampling of size $n = 2$ taken with repetition.

a) Compute the mean μ and variance σ^2 of this population.

b) How many 2-tuples are in sample space?

c) List all 2-tuples, which are possible to be the result of sampling.

d) State the distributions of following sample functions:

   (i) $Y = \sum_{i=1}^{2} X_i$
   Compute $E(Y)$ and $Var(Y)$.

   (ii) $\overline{X} = \frac{1}{n} \sum_{i=1}^{2} X_i$
   Compute $E(\overline{X})$ and $Var(\overline{X})$.

   (iii) $S'^2 = \frac{1}{n} \sum_{i=1}^{2} (X_i - \overline{X})^2$
   Compute $E(S'^2)$.

   (iv) $S^2 = \frac{1}{n-1} \sum_{i=1}^{2} (X_i - \overline{X})^2$
   Compute $E(S^2)$.

e) Proove by reference to results form d) following statements:

   (i) $\mathrm{E}(Y) = n \cdot \mu$ and $\mathrm{Var}(Y) = n \cdot \sigma^2$

   (ii) $\mathrm{E}(\overline{X}) = \mu$ and $\mathrm{Var}(\overline{X}) = \sigma^2/n$

   (iii) $\mathrm{E}(S'^2) = \sigma^2 \cdot (n-1)/n$

   (iv) $\mathrm{E}(S^2) = \sigma^2$

Assignment 6-1-e)
Proof of general forms

$X_1, X_2, \dots, X_n$ random sample from distribution with mean $\mu < \infty$ and variance $\sigma^2 < \infty$

$Y = \sum_{i=1}^n X_i$

$$\mathrm{E}[Y] = \mathrm{E}\left[\sum_{i=1}^n X_i\right] \overset{ind.}{=} \sum_{i=1}^n \mathrm{E}[X_i] \overset{i.d.}{=} n \cdot \mathrm{E}[X_1] = n \cdot \mu \tag{1}$$

$$\mathrm{Var}[Y] = \mathrm{Var}\left[\sum_{i=1}^n X_i\right] \overset{ind.}{=} \sum_{i=1}^n \mathrm{Var}[X_i] \overset{i.d.}{=} \sum_{i=1}^n \mathrm{Var}[X_1] = n \cdot \mathrm{Var}[X_1] = n \cdot \sigma^2 \tag{2}$$

$$\mathrm{E}[\bar{X}] = \mathrm{E}\left[\frac{\sum_{i=1}^n X_i}{n}\right] \overset{i.i.d.}{=} \frac{1}{n}\sum_{i=1}^n \mathrm{E}[X_1] = \frac{n\,\mathrm{E}[X_1]}{n} = \mu \tag{3}$$

$$\mathrm{Var}[\bar{X}] = \mathrm{Var}\left[\frac{\sum_{i=1}^n X_i}{n}\right] \overset{i.i.d.}{=} \frac{1}{n^2}\sum_{i=1}^n \mathrm{Var}[X_1] = \frac{n\,\mathrm{Var}[X_1]}{n^2} = \frac{\sigma^2}{n} \tag{4}$$

$$\begin{aligned}
\mathrm{E}[S^2] &= \frac{1}{n-1}\,\mathrm{E}\left[\sum_{i=1}^n \left(X_i - \mu + \mu - \bar{X}\right)^2\right] \\
&= \frac{1}{n-1}\,\mathrm{E}\left[\sum_{i=1}^n \left((X_i - \mu) - (\bar{X} - \mu)\right)^2\right] \\
&= \frac{1}{n-1}\,\mathrm{E}\left[\sum_{i=1}^n \left((X_i - \mu)^2 - 2(X_i - \mu)(\bar{X} - \mu) + (\bar{X} - \mu)^2\right)\right] \\
&= \frac{1}{n-1}\,\mathrm{E}[\sum_{i=1}^n (X_i - \mu)^2 - 2(\bar{X} - \mu)\underbrace{\sum_{i=1}^n (X_i - \mu)}_{n\bar{X} - n\mu} + n(\bar{X} - \mu)^2] \\
&= \frac{1}{n-1}\,\mathrm{E}\left[\sum_{i=1}^n (X_i - \mu)^2 - 2n(\bar{X} - \mu)^2 + n(\bar{X} - \mu)^2\right] \\
&= \frac{1}{n-1}\,\mathrm{E}\left[\sum_{i=1}^n (X_i - \mu)^2 - n(\bar{X} - \mu)^2\right] \\
&= \frac{1}{n-1}\left(\sum_{i=1}^n \underbrace{\mathrm{E}[(X_i - \mu)^2]}_{\mathrm{Var}(X) = \sigma^2} - n\underbrace{\mathrm{E}[(\bar{X} - \mu)^2]}_{\mathrm{Var}(\bar{X}) = \frac{\sigma^2}{n}}\right) \\
&= \frac{1}{n-1}(n\sigma^2 - \sigma^2) = \sigma^2
\end{aligned} \tag{5}$$

$$\mathrm{E}[S'^2] = \mathrm{E}\left[\frac{n-1}{n}S^2\right] = \frac{n-1}{n}\sigma^2 \tag{6}$$