Statistics I: Exercise Session 4

17.6.2015, 14-16, SPA1 21b

1 Descriptive statistics

Assume two samples of n observations $x_1, \ldots, x_n$ and $y_1, \ldots, y_n$.

Sample mean:	$\overline{x}$	$=$	$\frac{1}{n}\sum_{i=1}^{n} x_i$
Sample variance:	s_x^2	$=$	$\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2$
Empirical covariance:	s_{xy}	$=$	$\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})$

Bravais-Pearson correlation coefficient

$$r_{xy} = r_{yx} = \frac{s_{xy}}{s_x \cdot s_y} = \frac{\sum\limits_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum\limits_{i=1}^{n}(x_i - \overline{x})^2 \sum\limits_{i=1}^{n}(y_i - \overline{y})^2}}$$

$$= \frac{n\sum\limits_{i=1}^{n} x_i y_i - \sum\limits_{i=1}^{n} x_i \sum\limits_{i=1}^{n} y_i}{\sqrt{\left\{n\sum\limits_{i=1}^{n} x_i^2 - \left(\sum\limits_{i=1}^{n} x_i\right)^2\right\}\left\{n\sum\limits_{i=1}^{n} y_i^2 - \left(\sum\limits_{i=1}^{n} y_i\right)^2\right\}}}$$

with $-1 \leq r_{xy} \leq 1$

2 Combinatorics

	Without repetition	With repetition
Permutations	$P(n) = n!$	$P(n; g_1, \ldots, g_r) = \frac{n!}{g_1! \ldots g_r!}$
Variations	$V(n,k) = \frac{n!}{(n-k)!}$	$V^R(n,k) = n^k$
Combinations	$C(n,k) = \frac{n!}{k!(n-k)!} = \binom{n}{k}$	$C^R(n,k) = \binom{n+k-1}{k}$

Permutations	Any order of n elements
Variations	Selection of k of n elements <u>with</u> respect to order
Combinations	Selection of k of n elements <u>without</u> respect to order

3 Probability Theory

3.1 Events

$A = S$	A is a sure event
$A = \emptyset$	A is an impossible event
$A \subset B$	A is a subset of B
$A \equiv B$	A and B are equivalent events
$A \cap B = \emptyset$	A and B are disjoint events
$B = \overline{A}$	A and B are complementary events
$A = \bigcup_i A_i$	A is a union of events A_i
$A = \bigcap_i A_i$	A is an intersection of events A_i

Exercises

1-83: *Outdoor temperature and journey duration*

Student E measured an outdoor temperature X (in degrees Celsius) and a duration of his journey to university Y (in minutes):

x_i	-20	-10	0	10	20
y_i	60	40	35	20	20

How strong is the correlation between these two characteristics?

2-4: *Birthday party*

It is your birthday. However, you can only invite 6 of your 12 friends (all similar) to your party.

a) How many choices do you have to select the guests?

b) How many possible seating arrangements are there for your 6 guests at the birthday table?

c) You have invited 3 male friends and 3 female friends. How many possible seating arrangements do you have when 3 male friends and 3 female friends are considered in each case as the same?

2-14: *Trails*

Trails are marked with signs consisting of 2 coloured lines. How many colours are used

a) for 36 trails where the order of colours is considered and their repetition is accepted?

b) for 21 trails where the order of colours is not considered and their repetition is not allowed?

c) for 15 trails where the order of colours is not considered and their repetition is allowed?

3-1: *Dice*

The random experiment 'throw a dice twice' was done. Define the events: $A = \{6 \text{ in the first throw}\}$ and $B = \{6 \text{ in the second throw}\}$.

a) Set the sample space S for this experiment.

b) Calculate the number of elementary events of S using combinatorics.

c) Define the events A and B in the meaning of the elementary events.

d) Determine the union and the intersection of the two events A and B.

e) Set the sample space S, events A and B, their union and intersection in the Venn-diagram.

f) Give an impossible event of this random experiment.

g) Give the complementary events for A and B respectively.

h) Does $A \subset B$ hold?

i) Are the events A and B disjoint?

3-7: *Non-disjoint subsets*

A, B and C are non-disjoint subsets of sample space S. Using only the symbols of union, intersection, difference and complement of the event and letters A, B and C, write down the expressions of following events:

a) at least one of A, B and C occurs;

b) only A occurs;

c) A and B occur, but C does not;

d) all three occur;

e) none of A, B and C occurs;

f) exactly one of A, B and C occurs;

g) at most two of A, B and C occur.