Statistik I - Exercise session 6
7.7.2014 & 14.7.2014

Info

- Classroom: SPA1 220
- Time: Mondays, 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
28.04.14	E1	1-2, 1-3 (even), 1-10
05.05.14	E1	1-2, 1-3 (even), 1-10
12.05.14	E2	1-20, 1-22, 1-32
19.05.14	E2	1-20, 1-22, 1-32
26.05.14	E3	1-80, 1-83, (1-98)
02.06.14	E3	1-80, 1-83, (1-98)
09.06.14	–	–
16.06.14	E4	2-4, 2-14, 3-1, 3-7, (3-11)
23.06.14	E5	3-25, 3-37, 3-55
30.06.14	E5	3-25, 3-37, 3-55
07.07.14	E6	3-61,4-1,4-7, (4-29)
14.07.14	E6	3-61,4-1,4-7, (4-29)

Review

- week 11 & week 12
- Slides: random variables(1-45)

Conditional Probability $P(A|B) = \dfrac{P(A \cap B)}{P(B)}$, $P(B) > 0$ and $P(B|A) = \dfrac{P(A \cap B)}{P(A)}$, $P(A) > 0$

Independent events $P(A|B) = P(A|\overline{B}) = P(A)$ and $P(B|A) = P(B|\overline{A}) = P(B)$

Distribution of discrete random variable

Probability mass function for discrete r.v.

$$
\begin{aligned}
P(X = x_i) &= f(x_i) \qquad (i = 1, 2, \dots) \\
P(a \leq X \leq b) &= \sum_{a \leq x_i \leq b} P(X = x_i) = \sum_{a \leq x_i \leq b} f(x_i)
\end{aligned}
$$

Cumulative distr. function for discrete r.v.

$$
F(x) = P(X \leq x) = \sum_{x_i \leq x} f(x_i)
$$

Distribution of continuous random variable

Probability mass function for continuous r.v.

$$
P(a < X \leq b) = \int_a^b f(x)\,dx \text{ for all} a, b \text{ such that } a \leq b
$$

Cumulative distr. function for continuous r.v.

$$
F(x) = P(-\infty < X \leq x) = \int_{-\infty}^{x} f(t)\,dt
$$

Parameters of random variable

Expected value

discrete variable $E(X) = \mu_X = \sum_{i=1}^{k} x_i \cdot f(x_i)$

continuous variable $E(X) = \mu_X = \int_{-\infty}^{+\infty} x \cdot f(x)dx$

Properties of expected value:

$$
\begin{aligned}
E(a + b \cdot X) &= a + b \cdot E(X) \text{ (a, b constant)} \\
E(X \pm Y) &= E(X) \pm E(Y)
\end{aligned}
$$

Variance

discrete variable $Var(X) = \sigma_X^2 = \sum_{i=1}^{k}(x_i - \mu_X)^2 \cdot f(x_i) = \sum_{i=1}^{k} x_i^2 \cdot f(x_i) - \mu_X^2$

continuous variable $Var(X) = \sigma_X^2 = \int_{-\infty}^{+\infty} (x - \mu_X)^2 \cdot f(x)\, dx = \int_{-\infty}^{+\infty} x^2 \cdot f(x)\, dx - \mu_X^2$

Properties of variance:

$$
\begin{aligned}
Var(X) &= E[\{X - E(X)\}^2] = E(X^2) - [E(X)]^2 \\
Var(a + b \cdot X) &= b^2 \cdot Var(X) \text{ } (a, b \text{ konstant}) \\
Var(X \pm Y) &= Var(X) + Var(Y) \pm 2 \cdot Cov(X, Y)
\end{aligned}
$$

Linear combination of random variable

Linear combination	$Z_1 = a \cdot X + b \cdot Y$	$Z_2 = a \cdot X - b \cdot Y$
Expected value	$E(Z_1) = a \cdot E(X) + b \cdot E(Y)$	$E(Z_2) = a \cdot E(X) - b \cdot E(Y)$
Variance	$Var(Z_1) = a^2 \cdot Var(X) + b^2 \cdot Var(Y) + 2 \cdot a \cdot b \cdot Cov(X, Y)$	
	$Var(Z_2) = a^2 \cdot Var(X) + b^2 \cdot Var(Y) - 2 \cdot a \cdot b \cdot Cov(X, Y)$	

two discrete random variables

Joint distribution

P.m.f. $P(X = x_i, Y = y_j) = f(x_i, y_j)$

C.d.f. $F(x, y) = P(X \leq x, Y \leq y) = \sum_{x_i \leq x} \sum_{y_j \leq y} f(x_i, y_j)$

Marginal distribution

P.m.f. for X $f(x_i) = P(X = x_i) = \sum_{j=1}^{r} f(x_i, y_j)$

P.m.f. for Y $f(y_j) = P(Y = y_j) = \sum_{i=1}^{m} f(x_i, y_j)$

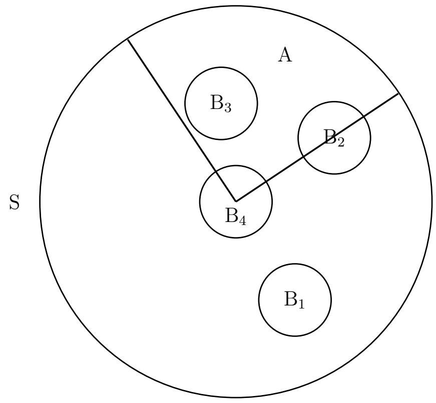
Marginal c.d.f. for X $P(X \leq x) = F(x) = \sum_{j=1}^{r} \sum_{x_i \leq x} f(x_i, y_j)$

Marginal c.d.f. for Y $P(Y \leq y) = F(y) = \sum_{y_j \leq y} \sum_{i=1}^{m} f(x_i, y_j)$

Exercises

Exercise 3-61 - Independent events

The probabilities of the events in the following figure are proportional to the respective areas:



Which one(s) of the following pairs of events consist(s) of independent events?

a) B_1, B_2 c) B_1, B_4 e) B_2, B_4 g) A, B_1 i) A, B_3

b) B_1, B_3 d) B_2, B_3 f) B_3, B_4 h) A, B_2 j) A, B_4

Exercise 4-1 - Children

The following table shows the number of children from six people:

Person	P_1	P_2	P_3	P_4	P_5	P_6
Number of children	6	2	0	1	2	1

From these people, three of them were selected. Let X be a total number of children from these three selected persons.

a) Specify a table of probability mass function

b) Specify a cumulative distribution function

c) Plot p.m.f. and c.d.f.

d) What is the probabilit that three randomly selected individuals have total number of children:

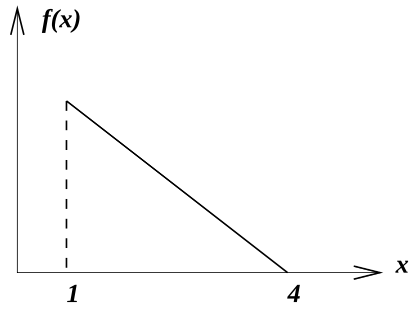
- less then or equal to 4

- more than 8

- more than 3 but less than 9 ?

Exercise 4-7 - Literature

The annual expenses of the student for literature are approximately triangular distributed, see figure:



a) Calculate the density function $f(x)$ and c.d.f $F(x)$.

b) Determine $\mathrm{E}(X)$ and $\mathrm{var}(X)$

c) Calculate the following probabilities: P(X ≤ 2), P($2 \leq X \leq 3$), P(X ≥ 3),

Exercise 4-29 - Two-dimensional random variable

Two-dimensional random variable (X_1, X_2) has the following probability mass function:

X_1	X_2		
	1	2	3
1	0,1	0,3	0,2
2	0,1	0,1	0,2

Determine the expected value of $X_3 = X_1 + X_2$

4-1-A

Person	P_1	P_2	P_3	P_4	P_5	P_6
# of kids	6	2	0	1	2	1

Σ kids	COMBINATIONS OF PERSONS		TOTAL POS COMBIN	frequency
2	(P_3, P_4, P_6)	(0,1,1)	1	1/20
3	(P_2, P_3, P_4) (P_2, P_3, P_6) (P_3, P_4, P_5) (P_3, P_5, P_6)	(2,0,1), (2,0,1) (0,1,2), (0,2,1)	4	4/20
4	(P_2, P_4, P_6) (P_2, P_3, P_5) (P_4, P_5, P_6)	(2,1,1) (2,0,2) (1,2,1)	3	3/20
5	(P_2, P_4, P_5) (P_2, P_5, P_6)	(2,1,2) (2,2,1)	2	2/20
7	(P_1, P_3, P_4) (P_1, P_3, P_6)	(6,0,1) (6,0,1)	2	2/20
8	(P_1, P_2, P_3) (P_1, P_3, P_5) (P_1, P_4, P_6)	(6,2,0) (6,0,2) (6,1,1)	3	3/20
9	(P_1, P_2, P_4) (P_1, P_2, P_6) (P_1, P_5, P_6) (P_1, P_4, P_5)	(6,2,1), (6,2,1) (6,2,1), (6,1,2)	4	4/20
10	(P_1, P_2, P_5)	(6,2,2)	1	1/10