Statistics I: Exercise Session 6

15.7.2015, 14-16, SPA1 21b

1 Univariate Statistics

1.1 Distribution of discrete random variables

Probability mass function of a discrete r. v.

$$f(x_i) = \mathrm{P}(X = x_i), \quad i = 1, 2, \ldots$$

Cumulative distribution function of a discrete r. v.

$$F(x) = \sum_{x_i \le x} f(x_i) = \mathrm{P}(X \le x), \quad i = 1, 2, \ldots$$

1.2 Distribution of continuous random variables

Probability mass function of a continuous r. v.

$$\int_a^b f(x_i)dx = \mathrm{P}(a \le X \le b), \quad \forall a, b, \text{ such that } a \le b$$

Cumulative distribution function of a continuous r. v.

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

1.3 Parameters of random variables

Expected value

Discrete r. v. $\mathrm{E}[X] = \mu_X = \sum_{i=1}^{k} x_i \cdot f(x_i)$

Continuous r. v. $\mathrm{E}[X] = \mu_X = \int_{-\infty}^{+\infty} x \cdot f(x)dx$

Properties of expected value

$\mathrm{E}[a + b \cdot X] \quad = a + b \cdot \mathrm{E}(X)$ (a, b constant)

$\mathrm{E}[X \pm Y] \quad = \mathrm{E}[X] \pm \mathrm{E}[Y]$

Variance

Discrete r. v. $\mathrm{Var}(X) = \sigma_X^2 = \sum_{i=1}^{k}(x_i - \mu_X)^2 \cdot f(x_i) = \sum_{i=1}^{k} x_i^2 \cdot f(x_i) - \mu_X^2$

Continuous r. v. $\mathrm{Var}(X) = \sigma_X^2 = \int_{-\infty}^{+\infty}(x - \mu_X)^2 f(x)dx = \int_{-\infty}^{+\infty} x^2 f(x)dx - \mu_X^2$

Generally $\mathrm{Var}(X) = \sigma_X^2 = \mathrm{E}[(X - \mathrm{E}[X])^2] = \mathrm{E}[X^2] - (\mathrm{E}[X])^2$

Properties of variance

$$\mathrm{Var}[a + b \cdot X] = b^2 \cdot \mathrm{Var}(X) \ (a,\ b \text{ constant})$$

$$\mathrm{Var}[X \pm Y] = \mathrm{Var}(X) + \mathrm{Var}(Y) \pm 2 \cdot \mathrm{Cov}(X,Y)$$

1 Bivariate Statistics

1.1 Two discrete random variables

Joint distribution

Probability mass function $\mathrm{P}(X = x_i, Y = y_j) = f(x_i, y_j), \quad i = 1, \ldots, m,\ j = 1, \ldots, r$

Cumulative distr. function $F(x, y) = \mathrm{P}(X \leq x, Y \leq y) = \sum_{x_i \leq x} \sum_{y_j \leq y} f(x_i, y_j)$

Marginal distribution

Pmf of X $\quad f(x_i) = \mathrm{P}(X = x_i) = \sum_{j=1}^{r} f(x_i, y_j)$

Pmf of Y $\quad f(y_j) = \mathrm{P}(Y = y_j) = \sum_{i=1}^{m} f(x_i, y_j)$

Cdf of X $\quad \mathrm{P}(X \leq x) = F(x) = \sum_{j=1}^{r} \sum_{x_i \leq x} f(x_i, y_j)$

Cdf of Y $\quad \mathrm{P}(Y \leq y) = F(y) = \sum_{i=1}^{m} \sum_{y_j \leq y} f(x_i, y_j)$

Exercises

4-1: *Children*

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

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

Assume that three of these people were selected. Let X be a total number of children of the selected individuals.

a) Specify a table of probability mass function of X.

b) Specify a cumulative distribution function of X.

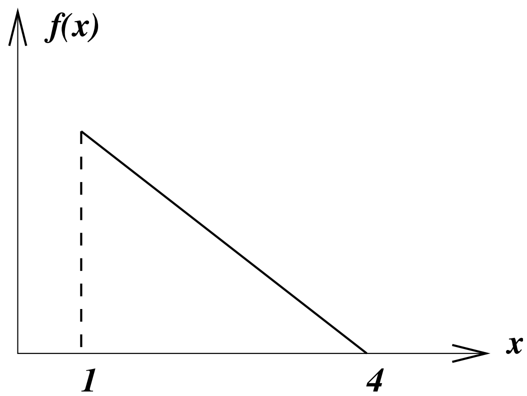
c) Plot the functions from a) and b).

d) What is the probability that the total number of children of three randomly selected people is

- less than or equal to 4?
- greater than 8?
- greater than 3 and less than 9?

4-7: *Literature*

Annual expenditure of a student on literature (X) is approximately triangular distributed (see figure below).



a) Calculate the pdf $f(x)$ and the cdf $F(x)$ of the expenditure.

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

c) Calculate the following probabilities: $\mathrm{P}(X \leq 2)$, $\mathrm{P}(2 \leq X \leq 3)$, $\mathrm{P}(X \geq 3)$.

4-29: *Two-dimensional random variable*

The 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 $X_3 = X_1 + X_2$.