

Assignment 1

Exercise 1

In a class of 23 students, what is the probability that at least two people have the same birthday?

Solution

Let's estimate the probability that no one matched birthdays at the group - $\bar{p}(n)$. We can do it by fixing the random person in the group. Then we have to take the second random person in the group. Probability of the fact that birthdays are not the same is $1 - \frac{1}{365}$. Then let us to take the third random person. Probability of the fact that their birthdays are not the same is $1 - \frac{2}{365}$. E.t.c.

$$\bar{p}(n) = \left(1 - \frac{1}{365}\right) \left(1 - \frac{2}{365}\right) \left(1 - \frac{3}{365}\right) \dots \left(1 - \frac{n-1}{365}\right) = \frac{365!}{365^n (365-n)!}$$

If there is 23 students at the group:

$$\bar{p}(n) = \frac{365!}{365^{23} (365-23)!} = 0.4927$$

The probability that at least two people from the group have the same birthday is:

$$p(n) = 1 - \bar{p}(n) = 0.5073$$

Answer: $p(n) = 0.5073$

Exercise 2

From a group of families with two children, one family is selected. Describe the space of elementary events. Assuming all elementary events equally probable, consider the random event A: there are a boy and a girl in that family, and the random event B: there is no more than one girl in the family. Calculate $P(A)$, $P(B)$, and $P(A \cap B)$. Are the events A and B independent?

Solution

We have the following space of events: two girls, two boys, the first child is a boy and the second is a girl, the first child is a girl, the second one is a boy.

- C - the probability that the child is boy is equal $\frac{1}{2}$.
- D - the probability that the child is girl is equal $\frac{1}{2}$.

The probability of A is:

$$P(A) = P(C|D) + P(D|C) = \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{2}$$

$$P(B) = \frac{3}{4} \text{ according to our sample space.}$$

If event A is occurred it means that there is one boy and one boy in the family. It B occurred always if A had occurred. Thus

$$P(B|A) = 1$$

Hence,

$$P(A \cap B) = P(B|A) \cdot P(A) = 1 \cdot \frac{1}{2} = \frac{1}{2}.$$

Answer: A and B are not independent by definition of independence.

Exercise 3

Three fair dice are rolled. What is the probability of obtaining at least one 6 if it is known that all the three dice showed different faces?

Solution

Let us define the requested probability p.

$$p = 1 - \bar{p} = 1 - \frac{5}{6} * \frac{4}{5} * \frac{3}{4} = \frac{1}{2}$$

$$\text{Answer: } p = \frac{1}{2}$$

Exercise 4

Suppose a breathalyzer has 5% false positives and 8% false negatives. That is, only 5% of the time will it indicate that a person is drunk when he is actually sober and 8% of the time will it indicate that a person is sober when the person is in fact drunk. Using this test, the police spot test a population of drivers, 99% of whom are sober. What is the chance that a person, who tests as drunk, is actually sober?

Solution

$p(D)$ - the probability the test shows that a person is is drunk:

$$p(D) = p(D|drunk) \cdot p(drunk) + p(D|sober) \cdot p(sober) = (1 - 0.08) \cdot 0.01 + 0.05 \cdot 0.99 = 0.0587$$

According to Bayes' rule:

$$p(D) = \frac{p(D|sober) \cdot p(sober)}{p(D)} = \frac{0.05 \cdot 0.99}{0.0587} = 0.84327$$

$$\text{Answer: } p(D) = 0.843$$

Exercise 5

Two fair dice are rolled. Let X_1 denote the number of points shown by the first die and X_2 denote the number of points shown by the second die. Consider the following events:

$$A_1 = \{X_1 \text{ is divisible by 2, } X_2 \text{ is divisible by 3}\}$$

$$A_2 = \{X_1 \text{ is divisible by 3, } X_2 \text{ is divisible by 2}\}$$

$$A_3 = \{X_1 \text{ is divisible by } X_2\}$$

$$A_4 = \{X_2 \text{ is divisible by } X_1\}$$

$$A_5 = \{X_1 + X_2 \text{ is divisible by 2}\}$$

$$A_6 = \{X_1 + X_2 \text{ is divisible by 3}\}$$

Solution

It is obvious that:

$$P(A_1) = \frac{1}{6} \quad P(A_2) = \frac{1}{6} \quad P(A_3) = \frac{7}{18} \quad P(A_4) = \frac{7}{18} \quad P(A_5) = \frac{1}{2} \quad P(A_6) = \frac{1}{3}$$

So let's find all jointly probabilities

$$\begin{aligned} P(A_1 \cap A_2) &= \frac{1}{36} & P(A_2 \cap A_3) &= \frac{1}{18} & P(A_3 \cap A_4) &= \frac{1}{6} & P(A_4 \cap A_5) &= \frac{1}{4} & P(A_5 \cap A_6) &= \frac{1}{6} \\ P(A_1 \cap A_3) &= \frac{1}{36} & P(A_2 \cap A_4) &= \frac{1}{18} & P(A_3 \cap A_5) &= \frac{1}{4} & P(A_4 \cap A_6) &= \frac{1}{6} \\ P(A_1 \cap A_4) &= \frac{1}{36} & P(A_2 \cap A_5) &= \frac{1}{36} & P(A_3 \cap A_6) &= \frac{1}{6} \\ P(A_1 \cap A_5) &= \frac{1}{36} & P(A_2 \cap A_6) &= \frac{1}{36} \\ P(A_1 \cap A_6) &= \frac{1}{36} \end{aligned}$$

Answer:

By definition of independence only $\{A_1 \text{ and } A_2\}$, $\{A_5 \text{ and } A_6\}$ are independent.

There is no reason to check triples and another combinations because it will include not independent events and so will not be independent.

Exercise 6

A) Let X be a random variable with a uniform distribution, i.e. with pdf equal to 1 for $x \in [0, 1]$ and 0 otherwise. Find the pdf of $Y = \exp(tX)$ for a fixed t . (This transformation is known as the moment generation function)

B) Find the pdf for this transformation ($\exp(t^*x)$) if the pdf of X is equal to $\exp(-x)$ for x positive and 0 otherwise.

Solution

According to the theorem from George Casella book

$$f_y(y) = f_x(g^{-1}(y)) * \frac{g^{-1}(y)}{dy} \text{ if } g(x) \text{ increases}$$

$$f_y(y) = -f_x(g^{-1}(y)) * \frac{g^{-1}(y)}{dy} \text{ if } g(x) \text{ decreases}$$

A)

$$f_x(x) = \begin{cases} 1, & x \in [0, 1] \\ 0 & \text{otherwise} \end{cases}$$

$$x \in [0, 1] \text{ so } y \in [1, e^t]$$

$$g(x) = e^{tx}$$

$$g^{-1}(y) = \frac{\ln(y)}{t}$$

$$\frac{g^{-1}(y)}{dy} = \frac{1}{ty}$$

$$f_y(y) = \frac{1}{ty}$$

B)

$$f_x(x) = \begin{cases} e^{-x}, & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

$$g(x) = e^{tx}$$

$$g^{-1}(y) = \frac{\ln(y)}{t}$$

$$\frac{g^{-1}(y)}{dy} = \frac{1}{ty}$$

$$f_y(y) = e^{-\frac{\ln(y)}{t}} * \frac{1}{ty}$$

Answer:

A)

$$f_y(y) = \begin{cases} \frac{1}{ty}, y \in [1, e^t] \\ 0 \text{ otherwise} \end{cases}$$

B)

$$f_y(y) = \begin{cases} y^{-\frac{1}{t}-1}, y > 0 \\ 0 \text{ otherwise} \end{cases}$$

Exercise 7

The random variable X has an exponential distribution: $F_x(x) = P(X \leq x) = 1 - e^{-x}$, ($x > 0$). What is the pdf f_x of X ? Find the probability density functions of the following random variables:

- a) $Y_1 = \sqrt{X}$
- b) $Y_2 = X^2$
- c) $Y_3 = \ln(X)$
- d) $Y_4 = \frac{1}{X}$
- e) $Y_5 = 1 - e^{-X}$
- f) Simulate 10000 draws of the random variable X : Use a histogram to show that the simulated X has the pdf $f_X(x)$
- g) Use the simulated X to generate 10000 draws of the random variable $Y_1 = \sqrt{X}$. Use a histogram for simulated Y_1 to show that your answer to (a) is correct.

Solution

$$F_x(X) = P(X \leq x) = 1 - e^{-x}, x \geq 0$$

$$f_x(x) = e^{-x}$$

a) $Y_1 = \sqrt{X}$

$$g^{-1}(y) = y^2$$

$$\frac{g^{-1}(y)}{dy} = 2y$$

$$f_y(y) = f_x(g^{-1}(y)) * \frac{g^{-1}(y)}{dy} = e^{-y^2} * 2y$$

$$f_y(y) = \begin{cases} e^{-y^2} * 2y, y > 0 \\ 0 \text{ otherwise} \end{cases}$$

b) $Y_2 = X^2$

$$F_y(y) = P(-\sqrt{y} \leq X \leq \sqrt{y}) = P(-\sqrt{y} < X \leq \sqrt{y}) = P(x \leq \sqrt{y}) - P(x \leq -\sqrt{y}) = F_x(\sqrt{y}) - F_x(-\sqrt{y}) = e^{\sqrt{y}} - e^{-\sqrt{y}}$$

$$f_y(y) = \frac{e^{\sqrt{y}} + e^{-\sqrt{y}}}{2\sqrt{y}}$$

$$f_y(y) = \begin{cases} \frac{e^{\sqrt{y}} + e^{-\sqrt{y}}}{2\sqrt{y}}, y \in (0, 0.231565] \\ 0 \text{ otherwise} \end{cases}$$

c) $Y_3 = \ln(X)$

$$g^{-1}(y) = e^y$$

$$\frac{g^{-1}(y)}{dy} = e^y$$

$$f_y(y) = f_x(g^{-1}(y)) * \frac{g^{-1}(y)}{dy} = e^{y-e^y}$$

d) $Y_4 = \frac{1}{X}$

$$g^{-1}(y) = \frac{1}{y}$$

$$\frac{g^{-1}(y)}{dy} = -\frac{1}{y^2}$$

$$f_y(y) = -f_x(g^{-1}(y)) * \frac{g^{-1}(y)}{dy} = \frac{e^{-\frac{1}{y}}}{y^2}$$

$$f_y(y) = \begin{cases} \frac{1}{y^2} e^{-\frac{1}{y}}, & y > 0 \\ 0 & \text{otherwise} \end{cases}$$

e) $Y_5 = 1 - e^{-X}$

$$g^{-1}(y) = -\ln(1 - y)$$

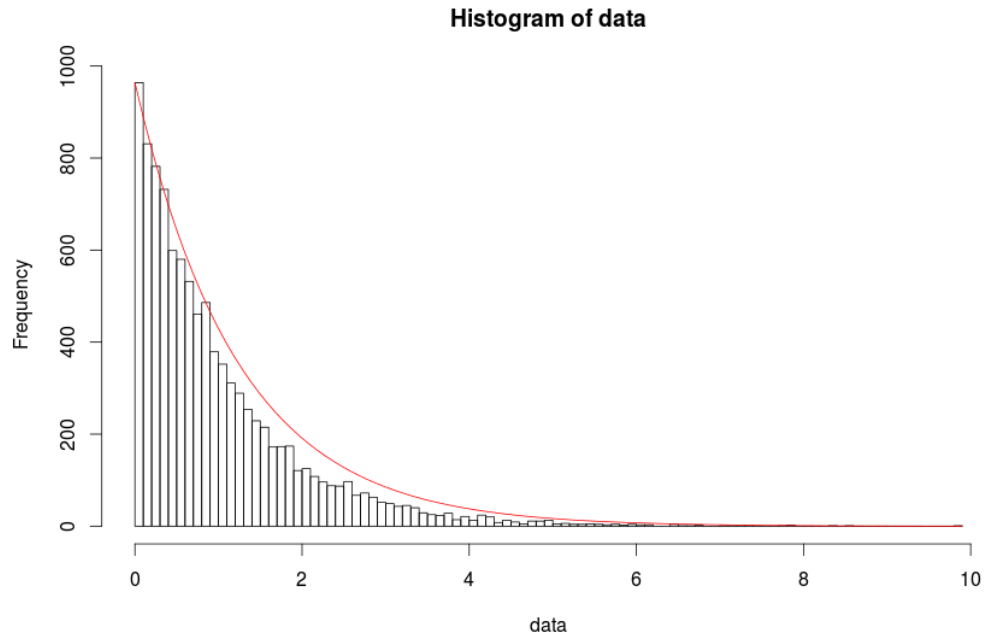
$$\frac{g^{-1}(y)}{dy} = \frac{1}{1 - y}$$

$$f_y(y) = -f_x(g^{-1}(y)) * \frac{g^{-1}(y)}{dy} = 1$$

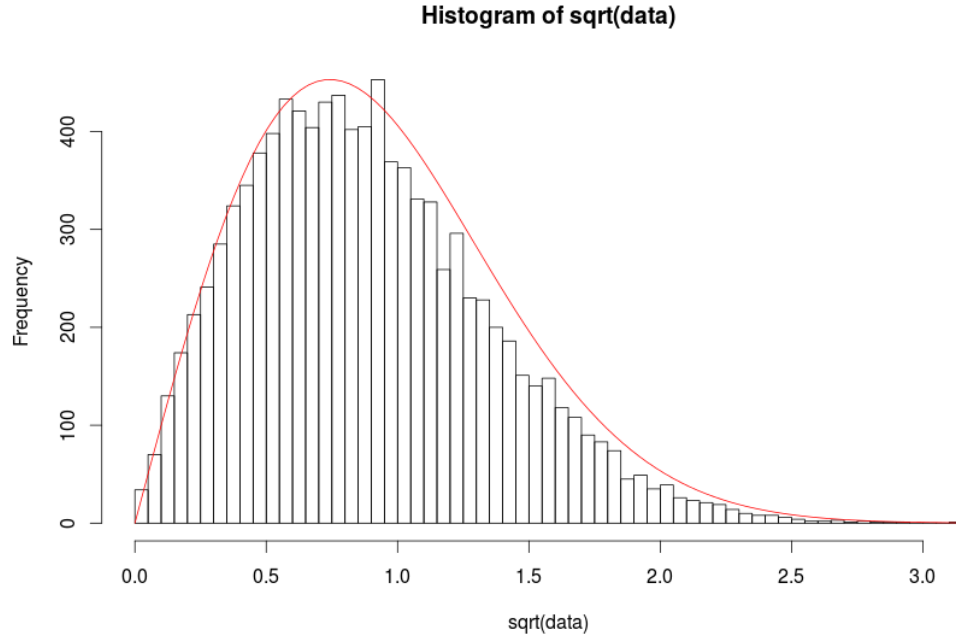
$$f_y(y) = \begin{cases} 1, & y \in [0, 1] \\ 0 & \text{otherwise} \end{cases}$$

f) Using Inverse transform sampling we can get random variable distributed by the given low, using uniform distributed variable for generation.

$$Y = -\ln(1 - X), X \in N[0, 1]$$



- g) Using Inverse transform sampling we can get random variable distributed by the given low, using uniform distributed variable for generation.



Answer:

Exercise 8

The pdf of a random vector (X, Y) is

$f_{X,Y}(x, y) = 24x^2y(1 - x)$, $0 < x < 1$, $0 < y < 1$ and $f_{X,Y}(x, y) = 0$ otherwise. Find

- The pdf and cdf of X ;
- The pdf and cdf of Y ;
- Are the random variables X and Y independent?

Solution

- $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy = \int_0^1 24x^2y(1 - x) dy = -12(1 - x)x^2$
- $f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx = \int_0^1 24x^2y(1 - x) dx = 2y$

- c) According to the Lemma 4.2.7 from "Statistical Inference" the X and Y are independent if:

$$f(x, y) = g(x)h(y)$$

Hence,

$$f_X(x)f_Y(y) = -12(1-x)x^2 \cdot 2y = -24x^2y(1-x)$$

$$f_X(x)f_Y(y) \neq f_{X,Y} - \text{X and Y are dependent}$$

Answer:

- a) $f_X(x) = -12(1-x)x^2$;
b) $f_Y(y) = 2y$;
c) X and Y are dependent.