Universität Bielefeld
Tanja Pasurek
Ehsan Abedi WiSe 2023/24

Foundations of Statistics

Solutions to Homework 6

Exercise 1 (Transformed density functions).

Let X be a continuous random variable. Define a new random variable $Y := g(X)$, where g is some map. In this exercise, we answer the following **important question:** What is the distribution of Y?

(a) Transformation formula in univariate case:

Let $X : \Omega \to \mathcal{I} \subseteq \mathbb{R}$ be a continuous random variable taking values in some interval $\mathcal{I}$. Suppose that the map $g : \mathcal{I} \to \mathbb{R}$ is strictly monotone increasing or strictly monotone decreasing, so that it has an inverse $h = g^{-1} : \mathcal{J} \to \mathcal{I}$ defined on the image set $\mathcal{J} := g(\mathcal{I}) = \{y \in \mathbb{R} : y = g(x) \text{ for some } x \in \mathcal{I}\}$. Moreover, assume that there exist a continuous derivative $g'(x) \neq 0$ for all $x \in \mathcal{I}$, which in turn guarantees the existence of

$$h'(y) = \frac{\partial}{\partial y} g^{-1}(y) = \frac{1}{g'(h(y))} \quad \text{for all } y \in \mathcal{J}.$$

Then show that the continuous random variable $Y := g(X)$ has PDF

$$f_Y(y) = \left|h'(y)\right| \cdot f_X(h(y)), \quad y \in \mathcal{J}. \tag{1}$$

Hint: use the so-called **CDF method**.

Apply formula (1) for the case of linear transformations and compare it with we found in Exercise 4(b), HW 4.

Solution: The map g is assumed to be strictly monotone (increasing or decreasing) and continuously differentiable with non zero-derivative. Then g is invertible and its inverse $h = g^{-1}$ is also strictly monotone (increasing or decreasing, respectively) and continuously differentiable with non zero-derivative.

- Let us first study the case that g is strictly *increasing*.

$$\begin{aligned} F_Y(y) &= \mathbb{P}(Y \leq y) \\ &= \mathbb{P}(g(X) \leq y) \\ &= \mathbb{P}(X \leq g^{-1}(y)) \\ &= F_X(g^{-1}(y)) \end{aligned}$$

Now take the derivative of $F_Y(y)$ w.r.t. y

$$\begin{aligned} f_Y(y) &= \frac{d}{dy} F_Y(y) \\ &= \frac{d}{dy} F_X(g^{-1}(y)) \\ &= F_X'(g^{-1}(y)) \frac{d}{dy} g^{-1}(y) \\ &= f_X(g^{-1}(y)) \frac{d}{dy} g^{-1}(y) \end{aligned}$$

• Now let us study the case that g is strictly *decreasing*.

$$\begin{aligned} F_Y(y) &= \mathbb{P}(Y \leq y) \\ &= \mathbb{P}(g(X) \leq y) \\ &= \mathbb{P}(X \geq g^{-1}(y)) \\ &= 1 - \mathbb{P}(X < g^{-1}(y)) \\ &= 1 - \mathbb{P}(X \leq g^{-1}(y)) \\ &= 1 - F_X(g^{-1}(y)) \end{aligned}$$

Now take the derivative of $F_Y(y)$ w.r.t. y

$$\begin{aligned} f_Y(y) &= \frac{d}{dy} F_Y(y) \\ &= -\frac{d}{dy} F_X(g^{-1}(y)) \\ &= -F_X'(g^{-1}(y)) \frac{d}{dy} g^{-1}(y) \\ &= -f_X(g^{-1}(y)) \frac{d}{dy} g^{-1}(y) \end{aligned}$$

• Considering both cases, we can write

$$f_Y(y) = \left| \frac{\partial}{\partial y} g^{-1}(y) \right| f_X(g^{-1}(y)), \quad y \in g(\mathcal{I}).$$

• Applying this formula to the linear transformation $x \mapsto g(x) = rx + s$ with $r > 0$ (which is indeed strictly monotone increasing and continuously differentiable), we first find $g^{-1}(y) = \frac{y-s}{r}$ and then obtain the CDF and PDF of the random variable $Y = rX + s$:

$$F_Y(y) = F_X(\tfrac{y-s}{r}), \qquad f_Y(y) = \tfrac{1}{r} f_X(\tfrac{y-s}{r}).$$

(b) Formula (1) can be generalized to multivariate case:

Let $\mathbf{X} : \Omega \to \boldsymbol{\mathcal{X}} \subseteq \mathbb{R}^n$ be a continuous random variable with joint density function $f_{\mathbf{X}}$. Let $g : \mathcal{X} \to \mathbb{R}^n$ be differentiable bijection with non-zero derivative. Prove that $\mathbf{Y} := g(\mathbf{X})$ has joint density function

$$f_{\mathbf{Y}}(\boldsymbol{y}) = \left|\det(J_{g^{-1}}(\boldsymbol{y}))\right| \cdot f_{\mathbf{X}}(g^{-1}(\boldsymbol{y})), \qquad \boldsymbol{y} \in \boldsymbol{\mathcal{Y}} := g(\boldsymbol{\mathcal{X}}), \tag{2}$$

whenever $J_{g^{-1}}(\boldsymbol{y})$ is well-defined. Here $J_{g^{-1}}(\boldsymbol{y})$ is the Jacobian matrix (i.e. the matrix of partial derivatives) of the map g^{-1}.

Hint: For each (Borel) subset $A \subset \mathbb{R}^n$, find $\mathbb{P}(\boldsymbol{Y} \in A)$ and apply the multi-dimensional **change of variables formula** as in *Calculus I*.

Solution: Let $A \subset \mathbb{R}^n$ be a (Borel) subset. Then

$$\begin{aligned}
\mathbb{P}(\boldsymbol{Y} \in A) &= \mathbb{P}(g(\boldsymbol{X}) \in A) \\
&= \mathbb{P}(\boldsymbol{X} \in g^{-1}(A)) \\
&= \int_{g^{-1}(A)} f_{\boldsymbol{X}}(\boldsymbol{x}) d\boldsymbol{x} \\
&= \int_A f_{\boldsymbol{X}}(g^{-1}(\boldsymbol{y})) |\det J_{g^{-1}}(\boldsymbol{y})| d\boldsymbol{y}
\end{aligned}$$

where $g^{-1}(A)$ is preimage of A under the map g and we used the multi-dimensional change of variables $\boldsymbol{x} = g^{-1}(\boldsymbol{y})$, $d\boldsymbol{x} = |\det J_{g^{-1}}(\boldsymbol{y})| d\boldsymbol{y}$. Note that g is differentiable bijection with non-zero derivative.

Equation above implies that the distribution of $\boldsymbol{Y}$ has density with respect to Lebesgue measure in $\mathbb{R}^n$ given by

$$f_{\boldsymbol{Y}}(\boldsymbol{y}) = \left|\det(J_{g^{-1}}(\boldsymbol{y}))\right| f_{\boldsymbol{X}}(g^{-1}(\boldsymbol{y})).$$

Exercise 2 (Transformed density functions, examples in 1D).

(a) Let a random variable X have PDF $f_X(x) = e^{-x}$ for $x > 0$. Define $Y := g(X) = \log X$. Using the above scheme, check that

$$f_Y(y) = e^y e^{-e^y} \quad \text{for } y \in \mathbb{R}.$$

(*Warning:* Although $Y := g(X)$, in general $f_Y \neq g(f_X)$!)

Solution: The map $x \mapsto g(x) := \log x$ is monotone increasing and continuously differentiable with non-zero derivative on $(0, \infty)$. We have $g^{-1}(y) = e^y$. Using (1), we obtain

$$f_Y(y) = |\tfrac{\partial}{\partial y} g^{-1}(y)| f_X(g^{-1}(y)) = e^y e^{-e^y} \qquad \forall y \in \mathbb{R}.$$

We note that this is *not* equal to $g(f_X(y)) = \log(e^{-y}) = -y$.

(b) Let $X \sim \mathrm{Unif}(0,1)$. Find the distribution of the random variable $Y = X^2$. Check your answer using simulation in `R`. To this end, simulate a large number (for instance, $n = 10^5$) of samples from the uniform distribution, square the values, make a histogram (with `freq=FALSE`) and superimpose the calculated density on top of the histogram. Compute $\mathbb{E}[Y]$ using both the LOTUS and the density f_Y that you have found.

Solution: The random variable Y takes values in $\mathcal{Y} = [0,1]$. The map $x \mapsto g(x) := x^2$ is monotone and continuously differentiable with non-zero derivative on $(0,1]$ and we have that $g^{-1}(y) = \sqrt{y}$. Using (1), we obtain

$$f_Y(y) = \left|\frac{\partial}{\partial y} g^{-1}(y)\right| f_X(g^{-1}(y)) = \frac{1}{2\sqrt{y}} \qquad \forall y \in (0,1],$$

and $f_Y(y) = 0$ otherwise.

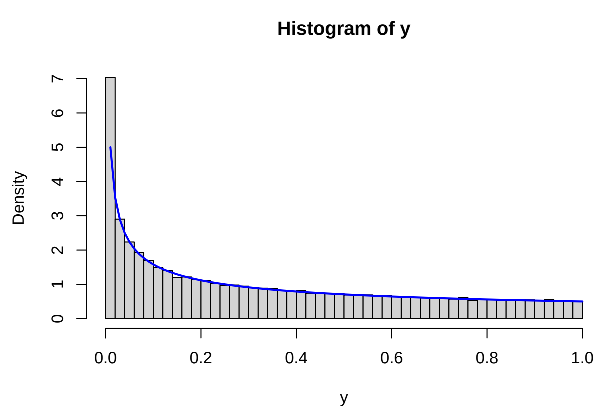
To compute expected value of Y, we first apply the so-called LOTUS (law of the unconscious statistician)

$$\mathbb{E}[Y] = \mathbb{E}[g(X)] = \int_0^1 g(x) f_X(x) dx = \int_0^1 x^2 dx = \frac{1}{3}.$$

Alternatively, we can use the density we found:

$$\mathbb{E}[Y] = \int_0^1 y f_Y(y) dy = \frac{1}{2}\int_0^1 \sqrt{y} dy = \frac{1}{2}\frac{2}{3} y^{3/2}\Big|_0^1 = \frac{1}{3}.$$

```r
x <- runif(n = 10**5, min = 0, max = 1)
y <- x^(2)
hist(y, br=50, freq=FALSE, xlab = 'y')
t <- seq(0.01, 1, by=0.01)
lines(t, 1/(2*sqrt(t)), col="blue", lwd=2)
```



Exercise 3 (Transformed density functions, examples in 2D).

(a) Let X and Y be independent, continuous random variables with densities f_X and f_Y. Use formula (2) to find the density of $X+Y$ and compare the result with what we found in Exercise 2, HW 5.

Solution: Let us define the map $g : \mathbb{R}^2 \to \mathbb{R}^2$ by

$$(x, y) \overset{g}{\mapsto} (x + y, y) \qquad (u, v) \overset{g^{-1}}{\mapsto} (u - v, v).$$

Then g is bijection and moreover, it has non-zero derivative:

$$J_{g^{-1}}(u, v) = \begin{bmatrix} \frac{\partial}{\partial u} g_1^{-1} & \frac{\partial}{\partial v} g_1^{-1} \\ \frac{\partial}{\partial u} g_2^{-1} & \frac{\partial}{\partial v} g_2^{-1} \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}$$

and $\det(J_{g^{-1}}) = 1$. Therefore

$$\begin{aligned} f_{X+Y,Y}(u, v) &= \left|\det(J_{g^{-1}}(u, v))\right| f_{X,Y}(g^{-1}(u, v)) \\ &= f_{X,Y}(u - v, v) \\ &= f_X(u - v) f_Y(v) \qquad \text{(by independence)} \end{aligned}$$

Finally integrate out v to find the first marginal:

$$\begin{aligned} f_{X+Y}(u) &= \int_{-\infty}^{\infty} f_{X+Y,Y}(u, v) dv \\ &= \int_{-\infty}^{\infty} f_X(u - v) f_Y(v) dv \end{aligned}$$

This is exactly the convolution formula we obtained in Ex. 2, HW 5.

(b) Let $X, Y \sim N(0, 1)$ be independent. Show that $X+Y$ is independent of $X - Y$. (*Hint:* define $U := X + Y$ and $V := X - Y$ and compute their joint density $f_{U,V}(u, v)$.)

Solution: First of all note that

$$U \sim N(0, 2), \quad V \sim N(0, 2),$$

which are the marginal distributions of the random vector $[U, V]^T$. To find their joint density function, let us define the map $g : \mathbb{R}^2 \to \mathbb{R}^2$ by

$$(x, y) \overset{g}{\mapsto} (x + y, x - y) \qquad (u, v) \overset{g^{-1}}{\mapsto} (\frac{u + v}{2}, \frac{u - v}{2})$$

Then g is bijection and moreover, it has non-zero derivative:

$$J_{g^{-1}}(u,v) = \begin{bmatrix} \frac{\partial}{\partial u} g_1^{-1} & \frac{\partial}{\partial v} g_1^{-1} \\ \frac{\partial}{\partial u} g_2^{-1} & \frac{\partial}{\partial v} g_2^{-1} \end{bmatrix} = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{bmatrix}$$

and $\det(J_{g^{-1}}) = -1/2$. Therefore

$$\begin{aligned}
f_{U,V}(u,v) &= \left|\det(J_{g^{-1}}(u,v))\right| f_{X,Y}(g^{-1}(u,v)) \\
&= \frac{1}{2} f_{X,Y}(\frac{u+v}{2}, \frac{u-v}{2}) \\
&= \frac{1}{2} f_X(\frac{u+v}{2}) f_Y(\frac{u-v}{2}) \\
&= \frac{1}{2} \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}(\frac{u+v}{2})^2) \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}(\frac{u-v}{2})^2) \\
&= \frac{1}{\sqrt{2}\sqrt{2\pi}} \exp(-\frac{1}{2}(\frac{u}{\sqrt{2}})^2) \frac{1}{\sqrt{2}\sqrt{2\pi}} \exp(-\frac{1}{2}(\frac{v}{\sqrt{2}})^2) \\
&= f_U(u) f_V(v)
\end{aligned}$$

We thus conclude that U and V are independent.

(c) Let $X_1, X_2 \sim N(0,1)$ be independent. Write the sample mean $\overline{X}$ and the sample variance S^2 in terms of X_1, X_2. Are $\overline{X}$ and S^2 independent or not?

Solution: We have

$$\begin{aligned}
\bar{X} &= \frac{X_1 + X_2}{2} \\
S^2 &= \frac{1}{n-1} \sum_i (X_i - \bar{X})^2 \\
&= \frac{1}{2-1} \{(\frac{X_1 - X_2}{2})^2 + (\frac{X_2 - X_1}{2})^2\} = \frac{1}{2}(X_1 - X_2)^2
\end{aligned}$$

Observe that

$$\begin{aligned}
\mathbb{P}(\bar{X} \le x, S^2 \le s) &= \mathbb{P}(\frac{X_1 + X_2}{2} \le x, \frac{1}{2}(X_1 - X_2)^2 \le s) \\
&= \mathbb{P}(X_1 + X_2 \le 2x, -\sqrt{2s} \le (X_1 - X_2) \le \sqrt{2s}) \\
&= \mathbb{P}(X_1 + X_2 \le 2x) \cdot \mathbb{P}(-\sqrt{2s} \le (X_1 - X_2) \le \sqrt{2s}) \qquad \text{(by part (b))} \\
&= \mathbb{P}(\bar{X} \le x) \cdot \mathbb{P}(S^2 \le s).
\end{aligned}$$

We thus conclude that $\bar{X}$ and S^2 are independent.

Exercise 4 (A universal random number generator, cf. Ch. 1.8).

Let $X : \Omega \to \mathbb{R}$ be a continuous random variable. Suppose that its CDF F_X is continuous and strictly increasing from 0 to 1 on some interval $\mathcal{I} \subseteq \mathbb{R}$. In this case, F_X has an inverse function $F_X^{-1} : [0, 1] \to \mathcal{I}$.

(a) Define $Y := F_X(X)$, i.e., you *plug a continuous random variable into its own CDF*. Show that $Y \sim \text{Unif}(0, 1)$. This is called the **probability integral transform**.

Solution: We first note that $Y \in [0, 1]$. We find the CDF of Y. Let $y \in [0, 1]$, then

$$\begin{aligned} F_Y(y) : &= \mathbb{P}(Y \leq y) \\ &= \mathbb{P}(F_X(X) \leq y) \\ &= \mathbb{P}(X \leq F_X^{-1}(y)) \\ &= F_X\big(F_X^{-1}(y)\big) = y. \end{aligned}$$

Therefore, we have

$$F_Y(y) = \begin{cases} 0 & y < 0 \\ y & 0 \leq y \leq 1 \\ 1 & y > 1 \end{cases}$$

Therefore, $Y \sim \text{Unif}(0, 1)$.

(b) Let now $U \sim \text{Unif}(0, 1)$ and define $Z := F_X^{-1}(U)$, i.e., you *plug a uniform random variable into an inverse CDF*. Show that Z and X have the same distribution, i.e., $F_Z = F_X$.

▶ **Conclusion:** *Any continuous real-valued random variable can be transformed into a uniform random variable and back by using its CDF.*

Solution: Let $z \in \mathbb{R}$. We have

$$\begin{aligned} F_Z(z) : &= \mathbb{P}(Z \leq z) \\ &= \mathbb{P}(F_X^{-1}(U) \leq z) \\ &= \mathbb{P}(U \leq F_X(z)) \\ &= F_X(z) \end{aligned}$$

(c) Write an `R` code to simulate continuous random variables from the density

$$f(x) = \frac{2}{(x+1)^3}, \quad x > 0.$$

Make a histogram of $n = 10^5$ simulated values and superimpose the density function to check the work.

Hint: The distribution is heavy-tailed, so in order to make a nice histogram, plot only the values less than 10 (which is about 99% of the values).

Solution: We apply the inverse transform method discussed in Ch. 1.8, which says that if F is a continuous and strictly increasing CDF and $U \sim \text{Unif}[0,1]$ is a uniformly distributed random variable, then the following random variable

$$X := F^{-1}(U)$$

has the cumulative distribution function F.
Let us first find the CDF corresponding to f

$$\begin{aligned} F(x) &= \int_0^x f(t)dt \\ &= \int_0^x \frac{2}{(t+1)^3} dt = 1 - \frac{1}{(x+1)^2}. \end{aligned}$$

Observe that F is continuous and strictly increasing. Write $y = F(x)$, then

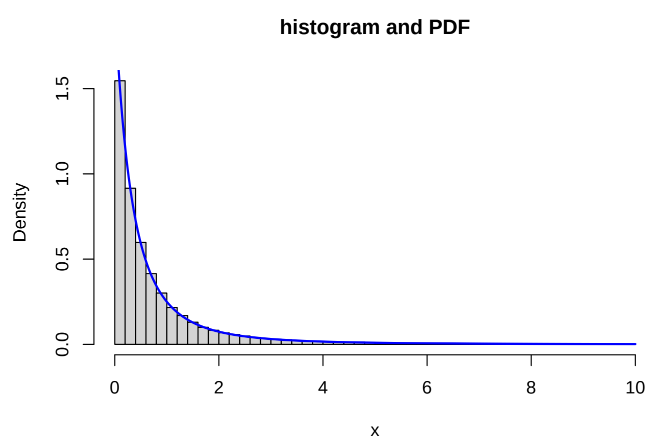
$$x = F^{-1}(y) = (1-y)^{-1/2} - 1.$$

and thus

$$X = (1-U)^{-1/2} - 1$$

has the cumulative distribution function F and density f.

```
u <- runif(10^5)
x <- (1-u)^(-1/2) - 1
hist(x[x < 10], br=50, freq=FALSE, xlab = 'x', main = 'histogram and PDF')
t <- seq(0, 10, by=0.01)
lines(t, 2/(t+1)^3, col="blue", lwd=2)
```



Exercise 5.

(a) Assume that $X : \Omega \to \mathbb{R}$ and $Y : \Omega \to \mathbb{R}$ are independent random variables, with X having a continuous distribution. Assume Y to be either discrete or continuous. Show that

$$\mathbb{P}\Big(\{\omega \in \Omega : X(\omega) \neq Y(\omega)\}\Big) = 1.$$

Solution:
• Let us first consider the case of Y being a discrete random variable. This means that Y takes values in a countable set (finite or infinite), which we denote by $u(i)$ with $i \in \mathcal{I}$ in a finite set or $i \in \mathbb{N}$, respectively. We prove for the case $i \in \mathbb{N}$ as the other case is similar,

$$\begin{aligned}
\mathbb{P}\big(\{X = Y\}\big) &= \mathbb{P}\Big(\{X = Y = u(1)\} \cup \{X = Y = u(2)\} \cup \cdots\Big) \\
&= \mathbb{P}\Big(\bigcup_{i \in \mathbb{N}} \{X = Y = u(i)\}\Big) \\
&= \sum_{i \in \mathbb{N}} \mathbb{P}\big(\{X = Y = u(i)\}\big) \\
&= \sum_{i \in \mathbb{N}} \mathbb{P}\big(\{X = u(i)\} \cap \{Y = u(i)\}\big) \\
&= \sum_{i \in \mathbb{N}} \underbrace{\mathbb{P}\big(X = u(i)\big)}_{=0} \mathbb{P}\big(Y = u(i)\big) \qquad \text{(by independence)} \\
&= 0.
\end{aligned}$$

• Second consider the case of Y being a continuous random variable with PDF $y \mapsto f_Y(y)$. Then $-Y$ is also a continuous random variable by (1) whose PDF is given $y \mapsto f_Y(-y)$. By HW 5, Exercise 2, we know that $Z := X - Y = X + (-Y)$ is also a continuous random variable (as a sum of two independent random variables). Then

$$\mathbb{P}\big(X = Y\big) = \mathbb{P}(Z = 0) = 0.$$

(b) Let X_n, $n \geq 1$, be a sequence of independent continuous random variables. Show that

$$\mathbb{P}\Big(\{\omega \in \Omega : X_i(\omega) = X_j(\omega) \text{ for some distinct indexes } i, j \geq 1\}\Big) = 0.$$

Solution: We have

$$
\begin{aligned}
&\mathbb{P}\big(\{X_i(\omega) = X_j(\omega) \text{ for some } i \neq j\}\big) \\
&\qquad\qquad\qquad = \mathbb{P}\Big(\bigcup_{i \neq j} \{X_i(\omega) = X_j(\omega)\}\Big) \\
&\qquad\qquad\qquad = \mathbb{P}\Big(\bigcup_{i=1}^{\infty} \bigcup_{j=i+1}^{\infty} \{X_i(\omega) = X_j(\omega)\}\Big) \\
&\qquad\qquad\qquad \leq \sum_{i=1}^{\infty} \sum_{j=i+1}^{\infty} \underbrace{\mathbb{P}\Big(\{X_i(\omega) = X_j(\omega)\}\Big)}_{=0 \text{ by (a)}} = 0
\end{aligned}
$$

(c) (From discrete to continuous uniform). Let X_n, $n \geq 1$, be a discrete random variable taking values in $\left\{\frac{1}{n+1}, ..., \frac{n}{n+1}\right\}$ uniformly. Show that X_n converges in distribution to the uniform distribution on $[0, 1]$.

Solution: Let us first recall that: the distribution of a sequence of real-valued random variables $(X_n)_{n \geq 1}$ with CDFs $(F_{X_n})_{n \geq 1}$ converges to the distribution with CDF F, if

$$
\lim_{n \to \infty} F_{X_n}(x) = F(x) \tag{3}
$$

for all $x \in \mathbb{R}$ *at which F is continuous.*

In this exercise, we want to show that CDFs converges to the CDF of a uniform distribution on $[0, 1]$, i.e.,

$$
F(x) = \begin{cases} 0 & \text{if } \quad x < 0 \\ x & \text{if } 0 \leq x < 1 \\ 1 & \text{if } 1 \leq x \end{cases}
$$

F is continuous at all $x \in \mathbb{R}$. Thus, we need to very (3) at *all* $x \in \mathbb{R}$.

We need to find the CDF of $F_{X_n}(x)$ for all values of $x \in \mathbb{R}$. Since $X_n \in [\frac{1}{n+1}, \frac{n}{n+1}]$, it is clear that $F_{X_n}(x) = 0$ for $x \in (-\infty, \frac{1}{n+1})$ and $F_{X_n}(x) = 1$ for $x \in [\frac{n}{n+1}, \infty)$. So it remains to consider x to be in the range

$$
\frac{1}{n+1} \leq x < \frac{n}{n+1}.
$$

Then there exists a natural number $1 \leq k < n$ such that

$$
\frac{k}{n+1} \leq x < \frac{k+1}{n+1}.
$$

In other words, $k = [x \cdot (n+1)]$. The random variable X_n takes n distinct values uniformly. Then probability of taking each is equal to $1/n$. Therefore, the CDF of X_n is a step function given by

$$F_{X_n}(x) = \begin{cases} 0 & \text{if } \quad x < \frac{1}{n+1} \\ \frac{k}{n} & \text{if } \frac{k}{n+1} \le x < \frac{k+1}{n+1},\ k = 1, 2, \cdots, n-1 \\ 1 & \text{if } \frac{n}{n+1} \le x \end{cases}$$

In the middle case above, we have

$$-\frac{x}{n} \le x - F_{X_n}(x) \le \frac{1}{n}(1-x).$$

Taking limit, we obtain

$$\lim_{n\to\infty} F_{X_n}(x) = \begin{cases} 0 & \text{if } \quad x < 0 \\ x & \text{if } 0 \le x < 1 \\ 1 & \text{if } 1 \le x \end{cases}$$

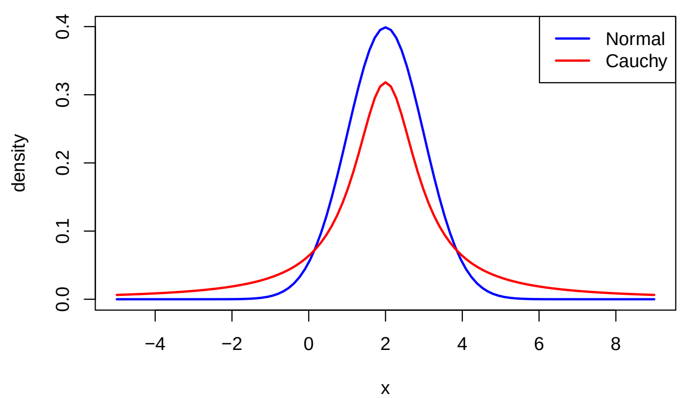
which is equal to F, the CDF of uniform distribution on $[0, 1]$.

Exercise 6 (Simulation of Law of Large Numbers (LLN) in R).

(a) To begin with, plot the densities of normal distribution with mean 2 and variance 1 (in blue) and Cauchy distribution with location parameter 2 and scale parameter 1 (in red) on the same plot. Which one has a heavier tail?

Solution:

```
# part (a)
curve(dnorm(x, mean = 2, sd = sqrt(1)),
      from = -5, to = 9, col = "blue", lwd = 2, ylab = "density", xlab = "x")
curve(dcauchy(x, location = 2, scale = 1),
      from = -5, to = 9, col = "red", lwd = 2, add = TRUE)
legend("topright", legend = c("Normal", "Cauchy"), col = c("blue", "red"), lwd = 2)
```




```
# we observe that Cauchy has heavier tail
```

(b) Take a sample of $n = 5000$ realizations from $N(2, 1)$. Calculate the cumulative arithmetic mean of your sample, that is the arithmetic mean of the first number, of the first two numbers, and so on (see ?`cumsum`). Plot the mean values obtained and overlaid them with a horisontal line corresponding to the actual mean value.

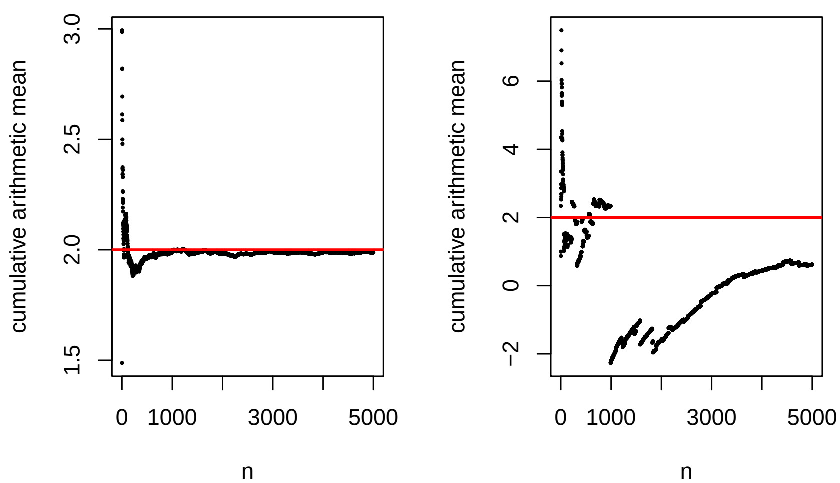
(c) Repeat the procedure in (b) with the Cauchy distribution with with location parameter 2 and scale parameter 1 (see ?`rcauchy`). Can we observe a similar convergence in this case? Justify your answer.

Hint: if you want to get a reproducible sequence of random numbers, use the command `set.seed` to start a random generator with any number of your choice (see e.g. [*Heumann et al.*], Appendix: Introduction to `R`, p. 418).

Solution:

```r
# part (b)
set.seed(22)
cum_sum <- cumsum(rnorm(n=5000,mean=2,sd=1))
cum_mean <- cum_sum/seq(1,5000,1)
par(mfrow=c(1,2))
plot(cum_mean,xlab='n',ylab='cumulative arithmetic mean',
     type='p',pch = 20, cex = 0.5, lwd=0.5)
abline(h=2,col='red',lwd=2)

# part (c)
cum_sum_c <- cumsum(rcauchy(n=5000,location =2,scale=1))
cum_mean_c <- cum_sum_c/seq(1,5000,1)
plot(cum_mean_c,xlab='n',ylab='cumulative arithmetic mean',
     type='p',pch = 20, cex = 0.5, lwd=0.5)
abline(h=2,col='red',lwd=2)
```



Remark: Try with different values for `set.seed()`. You will see that there is no clear convergence. The expected value of Cauchy distribution is undefined (in fact, all moments are undefined!). Therefore, it does not satisfy the assumption required for the Law of Large Numbers.

Exercise 7 (Simulation of Central Limit Theorem (CLT) in `R`).

Let $X_i, i \geq 1$ be a sequence of i.i.d. random variables with $\mathbb{E}[X_i] = \mu$ and $\mathrm{Var}(X_i) = \sigma^2$ such that $\sigma^2 \in (0, \infty)$. CLT tells us that the distribution of standardized sum

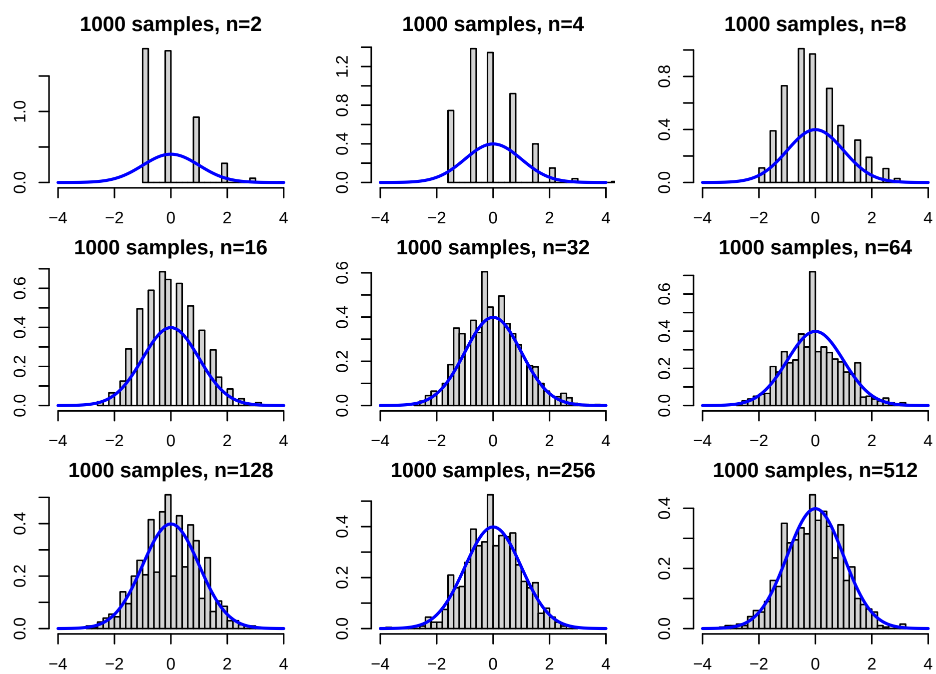
$$Z_n := \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} = \frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\sigma},$$

converges to the standard normal distribution $N(0, 1)$.

To see this, you need to consider two natural numbers k and n. Fix $k = 1000$. At first, take an arbitrary n and generate $k = 1000$ random samples Z_n when we have i.i.d. Pois(0.5)–distributed random variables X_i, $1 \leq i \leq n$. Plot the corresponding histogram and overlaid it with the density of the normal distribution. Then increase n, while keep k fixed. Repeat the simulation when we have i.i.d. Exp(2)–distributed random variables X_i, $1 \leq i \leq n$. Does the result depend on distribution of X_i?

Solution: see the results in next page. Both Poisson and Exponential distribution have finite variance. Therefore, CLT is applicable. We observe that the final result (as n becomes very large) does not depend on distribution of X_i and in both case the distribution of standardized sum Z_n converges to the standard normal distribution $N(0, 1)$. However, the *rate of convergence* depends on distribution of X_i.

```r
K = 1000
lambda = 0.5
rp <-function(n){
  Zn <- rep(NA, K)
  for (k in 1:K){
    s <- rpois(n, lambda = lambda)
    Zn[k] <- (sum(s) - n*lambda) / (sqrt(n)*sqrt(lambda))
    }
  return(Zn) # length K = 1000
  }
# Plot histogram of K samples of standardized sum with normal distribution curve over it
plot_hist<-function(n){
  Zn <- rp(n)
  hist(Zn, breaks=30, prob=T, xlim=c(-4,4), main=sprintf("%s samples, n=%s", K, n))
  curve(dnorm(x), col="blue", lwd=2, add=T, yaxt="n")
  }
par(mfrow=c(3,3), mar=c(2,2,2,2))
for (i in 1:9) {
  n <- 2^i
  plot_hist(n)
}
```



```r
K = 1000
lambda = 2
rp <-function(n){
  Zn <- rep(NA, K)
  for (k in 1:K){
    s <- rexp(n, rate = lambda)
    Zn[k] <- (sum(s) - n*1/lambda) / (sqrt(n)*sqrt(1/lambda^2))
  }
  return(Zn) # length K = 1000
}
# Plot histogram of K samples of standardized sum with normal distribution curve over it
plot_hist<-function(n){
  Zn <- rp(n)
  hist(Zn, breaks=30, prob=T, xlim=c(-4,4), main=sprintf("%s samples, n=%s", K, n))
  curve(dnorm(x), col="blue", lwd=2, add=T, yaxt="n")
}
par(mfrow=c(3,3), mar=c(2,2,2,2))
for (i in 1:9) {
  n <- 2^i
  plot_hist(n)
}
```