1 Random Variables Part II — Change of Variables

1.1 Change of Variables Formula (Discrete Random Variables)

Suppose that we know the PDF $f_X(x)$ of X. Our goal is to recover the the PMF $f_Y(y)$ of the random variable $Y = g(X)$. This can be done directly using the following steps

1. Using the range of X, find the range of $Y = g(X)$ denoted by computing the image of the range of X:
$$Y(S) = g(X(S)).$$
2. Compute the PMF of Y by expressing it in terms of the PMF of X:
$$f_Y(y) = \mathbb{P}(Y = y) = \mathbb{P}(g(X) = y) = \sum_{x:g(x)=y} \mathbb{P}(X = x) = \sum_{x\in g^{-1}(\{y\})=y} f_X(x).$$

This procedure is summarized by the change of variables formula.

Theorem 1 (*Change of Variables Formula (Discrete)*)

If X is discrete and $g : \mathbb{R} \to \mathbb{R}$, then
$$f_Y(y) = \sum_{x\in g^{-1}(\{y\})} f_X(x), \quad y \in Y(S) = g(X(S)).$$

1.2 Change of Variables Formula (Continuous Random Variables)

Suppose that we know the PDF $f_X(x)$ of X. Our goal is to recover the the PDF $f_Y(y)$ of the random variable $Y = g(X)$. This can be done directly using the following steps

1. Use the support of X to find the support of $Y = g(X)$:
$$\mathrm{supp}(Y) = \mathrm{cl}(\{y \in \mathbb{R} : f_Y(y) > 0\}) = g(\mathrm{supp}(X)).$$
2. Compute the CDF of Y for $y \in \mathrm{supp}(Y)$ by expressing it in terms of the CDF of X:
$$F_Y(y) = \mathbb{P}(g(X) \le y) = \dots$$
When the function g is not strictly increasing (or decreasing) over the support of X, then we must be careful when rewriting the inequality $\mathbb{P}(g(X) \le y)$.
3. Compute the PDF of Y by differentiating the CDF of Y,
$$f_Y(y) = F_Y'(y) \qquad y \in \mathrm{supp}(Y).$$

Remark 1. Technically, the function F_Y is only differentiable on the interior of $\mathrm{supp}(Y)$. This is not an issue since we can simply define $f_Y(y)$ on the boundaries by continuity.

When g is invertible, the above procedure gives us the change of density formula.

Theorem 2 (*Change of Variables Formula (Continuous)*)

Let X be a (absolutely) continuous random variable and g be invertible and differentiable with inverse g^{-1} on the support of Y, then
$$f_Y(y) = |(g^{-1})'(y)| f_X(g^{-1}(y)) = \frac{1}{|g'(g^{-1}(y))|} f_X(g^{-1}(y)), \quad y \in supp(Y).$$

1.3 Quantiles

Suppose that we are given X and a value $p \in (0, 1)$ and we are interested in computing the value of t such that

$$F_X(t) = \mathbb{P}(X \le t) = p.$$

If F_X is invertible, then $t = F_X^{-1}(p)$. However, not all CDFs are invertible so how does one define such a t in general. This generalized notation of an inverse is called a quantile function.

Definition 1. Let $p \in [0, 1]$. The p quantile (or $100 \times p$th percentile) of the distribution of X with CDF F_X is the smallest number c_p that satisfies $F_X(c_p) \ge p$. In other words,

$$c_p = \inf\{x \in \mathbb{R} : F_X(x) \ge p\}.$$

Remark 2. Recall that the infimum of a set A is the largest lower bound of A. For example,

$$\inf\{x \in \mathbb{R} : 0 < x < 1\} = 0.$$

Definition 2. The *median* of a distribution is its 0.5 quantile.

1.3.1 Generalized Inverse Interpretation

The quantile function $p \mapsto c_p$ is also called *generalized inverse function*, because it is a well defined function even if F_X is not strictly increasing like in the case of discrete random variables. It is an abuse of notation, but we often use the notation

$$F_X^{-1}(p) := c_p = \inf\{x \in \mathbb{R} : F_X(x) \ge p\}.$$

The quantile function $F_X^{-1}(p)$ is non-decreasing and left-continuous for $p \in (0, 1)$. We can compute the quantile in the following way

- If the distribution function F_X is continuous and strictly increasing, it has an inverse F_X^{-1} so

$$c_p = F_X^{-1}(p).$$

- If F_X has jumps or flat regions, then $F_X(x) = p$ may not have any solution or it might have infinitely many. In this case, the function $F_X^{-1}(p)$ is the left continuous step function that interpolates between the points (p, x) where x is the location of the jumps of F_X.

1.4 Example Problems

1.4.1 Applications

Problem 1.1. Let X be a continuous random variable with CDF $F_X(x) = \mathbb{P}(X \le x)$ and let $g : \mathbb{R} \mapsto \mathbb{R}$ be an increasing function with inverse g^{-1}. Compute $F_Y(y) = \mathbb{P}(Y \le y)$.

Solution 1.1. We can write the CDF of Y in terms of the CDF of X by

$$F_Y(y) = \mathbb{P}(g(X) \le y) = \mathbb{P}(X \le g^{-1}(y)) = F_X(g^{-1}(y)).$$

Remark 3. Notice that the CDF of $Y = g(X)$ is not $g(F_X(x))$.

Problem 1.2. Let X be a continuous random variable with the following pdf:

$$f_X(x) = \begin{cases} \frac{1}{4} & 0 < x \leq 4, \\ 0 & \text{otherwise} \end{cases}$$

1. Find the CDF of X.
2. Let $Y = X^{-1}$. Find the CDF Y.
3. Find the PDF of Y.

Solution 1.2.

1. Outside the support of X, we have $F_X(x) = 0$ for $x < 0$ and $F_X(x) = 1$ for $x > 4$. For x in the support of X, i.e. $x \in [0, 4]$
$$F_X(x) = \int_0^x \frac{1}{4} = \frac{x}{4}.$$
In summary,
$$F_X(x) = \begin{cases} 0, & \text{if } x < 0 \\ \frac{x}{4}, & \text{if } 0 \leq x < 4 \\ 1, & \text{if } 4 \leq x \end{cases}$$

2. Notice that $\text{supp}(X) = [0, 4]$. Therefore, the support of $Y = X^{-1}$ is
$$0 \leq X \leq 4 \implies \infty \geq X^{-1} \geq \frac{1}{4} \implies \text{supp}(Y) = \Big[\frac{1}{4}, \infty\Big).$$
For y in the support of Y, i.e. $y \in \Big[\frac{1}{4}, \infty\Big)$
$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(X^{-1} \leq y) = \mathbb{P}(X > \frac{1}{y}) = 1 - \mathbb{P}(X \leq \frac{1}{y}) = 1 - F_X(y^{-1}) = 1 - \frac{1}{4y}.$$
In summary,
$$F_Y(y) = \begin{cases} 0, & \text{if } y < \frac{1}{4} \\ 1 - \frac{1}{4y}, & \text{if } \frac{1}{4} \leq y \end{cases}$$

3. To get the PDF of Y, we simply differentiate $F_Y(y)$ to conclude that
$$f_Y(y) = F_Y'(y) = \frac{d}{dy}\Big(1 - \frac{1}{4y}\Big) = \frac{1}{4y^2} \qquad \text{for } y \in \Big[\frac{1}{4}, \infty\Big).$$
and $f_Y(y) = 0$ for $y \notin \Big[\frac{1}{4}, \infty\Big)$.

Alternative Solution: We can compute the PDF of f using the change of variables formula. We have $g(x) = \frac{1}{x}$, so $g'(x) = -\frac{1}{x^2}$ and $g^{-1}(x) = \frac{1}{x}$. Therefore, for $y \in [\frac{1}{4}, \infty)$,

$$f_Y(y) = \frac{1}{|g'(g^{-1}(y))|} f_X(g^{-1}(y)) = \frac{1}{|-\frac{1}{(y^{-1})^2}|}\frac{1}{4} = \frac{1}{4y^2}.$$

Problem 1.3. Suppose a continuous random variable X has probability density function

$$f_X(x) = \begin{cases} 1 - |x| & -1 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

1. Find the CDF of X.
2. Let $Y = X^2$. Find the CDF Y.
3. Find the PDF of Y.

Solution 1.3.

1. We can rewrite the PDF as
$$f_X(x) = \begin{cases} 1+x & -1 \le x < 0 \\ 1-x & 0 \le x < 1 \\ 0 & \text{otherwise} \end{cases}$$
Outside the support of X, we have $F_X(x) = 0$ for $x < -1$ and $F_X(x) = 1$ for $x > 1$. We have two cases for x in the support of X,
   (a) If $x \in [-1, 0]$
$$F_X(x) = \int_{-1}^{x} 1 + t \, dt = t + \frac{t^2}{2}\Big|_{-1}^{x} = x + \frac{x^2}{2} + \frac{1}{2}.$$
   (b) If $x \in [0, 1]$
$$F_X(x) = \int_{-1}^{0} 1 + t \, dt + \int_{0}^{x} 1 - t \, dt = \Big(t + \frac{t^2}{2}\Big)\Big|_{-1}^{0} + \Big(t - \frac{t^2}{2}\Big)\Big|_{0}^{x} = x - \frac{x^2}{2} + \frac{1}{2}.$$
In summary,
$$F_X(x) = \begin{cases} 0 & x < -1 \\ x + \frac{1}{2}x^2 + \frac{1}{2} & -1 \le x < 0 \\ x - \frac{1}{2}x^2 + \frac{1}{2} & 0 \le x < 1 \\ 1 & 1 \le x \end{cases}$$

2. Notice that $\operatorname{supp}(X) = [-1, 1]$. Therefore, the support of $Y = X^2$ is
$$-1 \le X \le 1 \implies 0 \le X^2 \le 1 \implies \operatorname{supp}(Y) = [0, 1].$$
Outside of the support, for $y < 0$ we have $F_Y(y) = 0$ and for $y > 1$ we have $F_Y(y) = 1$. In the support of Y, we have
$$\begin{aligned} F_Y(y) &= \mathbb{P}(Y \le y) = \mathbb{P}(X^2 \le y) = \mathbb{P}(-\sqrt{y} \le X \le \sqrt{y}) \\ &= F_X(\sqrt{y}) - F_X(-\sqrt{y}) \\ &= \Big(\sqrt{y} - \frac{1}{2}y + \frac{1}{2}\Big) - \Big(-\sqrt{y} + \frac{1}{2}y + \frac{1}{2}\Big) \\ &= 2\sqrt{y} - y. \end{aligned}$$
In summary,
$$F_Y(y) = \begin{cases} 0 & y < 0 \\ 2\sqrt{y} - y & 0 \le y < 1 \\ 1 & 1 \le y \end{cases}$$

3. To get the PDF of Y, we simply differentiate $F_Y(y)$ to conclude that

$$f_Y(y) = F_Y'(y) = \frac{d}{dy}\Big(2\sqrt{y} - y\Big) = \frac{1}{\sqrt{y}} - 1 \qquad \text{for } y \in [0,1].$$

and $f_Y(y) = 0$ for $y \notin [0,1]$.

Remark 4. The function $g(x) = x^2$ is not increasing on the $\text{supp}(X) = [-1,1]$ so we can't use the change of variables formula.

Problem 1.4. Suppose X has PDF

$$f_X(x) = \begin{cases} 2e^{-2x}, & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

What is the median of the distribution of X?

Solution 1.4.

- The cdf of X is

$$F_X(x) = \int_{-\infty}^{x} f(t)\,dt = \int_0^x 2e^{-2t}\,dt = -e^{-2t}\Big|_0^x = 1 - e^{-2x}, \quad x \geq 0$$

and $F_X(x) = 0$ for $x < 0$.

- The function F_X is strictly increasing for $x \geq 0$ so $F_X(c_p) = p$ has a unique solution for $p \in (0,1)$:

$$F_X(c_p) = p \Leftrightarrow 1 - e^{-2c_p} = p \Leftrightarrow c_p = -\log(1-p)/2$$

- For $p = 0.5$ we get $c_{0.5} = -\log(0.5)/2 \approx 3.466$ as the median.

Problem 1.5. Suppose X and Y are continuous random variables satisfying $\mathbb{P}(X \leq t) < \mathbb{P}(Y \leq t)$ for all $t \in \mathbb{R}$. Let s_x and s_y denote the median of the distributions of X and Y, respectively. Show that

$$s_y < s_x.$$

Solution 1.5. By definition, we have $0.5 = F_X(s_x) = F_Y(s_y)$. However, since $\mathbb{P}(X \leq t) < \mathbb{P}(Y \leq t)$

$$0.5 = F_Y(s_y) = F_X(s_x) < F_Y(s_x)$$

This implies that $F_Y(s_y) < F_Y(s_x)$ which must mean that $s_y < s_x$ because F_Y is increassing.

Problem 1.6. Consider the random variable X with

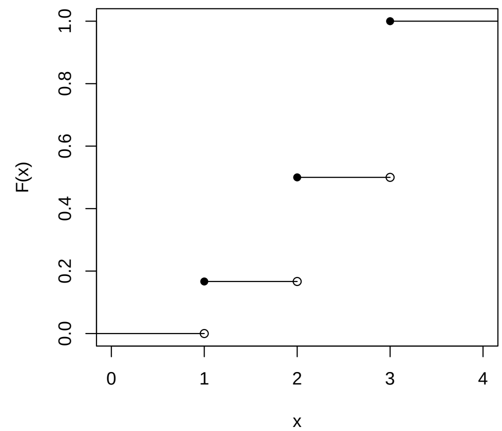
$$\mathbb{P}(X = 1) = 1/6, \qquad \mathbb{P}(X = 2) = 2/6 \qquad \mathbb{P}(X = 3) = 3/6.$$

Sketch the CDF of X and compute $F_X^{-1}(p)$ for $p \in (0,1)$.

Solution 1.6.

1. The cumulative distribution function is

$$F_X(x) = \begin{cases} 0, & x < 1, \\ \frac{1}{6}, & 1 \le x < 2 \\ \frac{1}{2}, & 2 \le x < 3, \\ 1 & 3 \le x \end{cases}$$

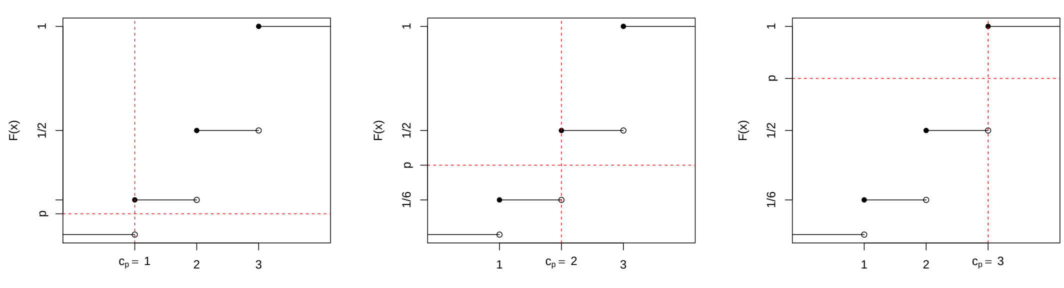


2. To compute the quantile function, we notice that the discontinuities of the CDF occur at $(1, \frac{1}{6}), (2, \frac{1}{2}), (3, 1)$. Therefore, the discontinuities for the quantile function occur at $(\frac{1}{6}, 1), (\frac{1}{2}, 2), (1, 3)$. Extending this to make the function left continuous implies

$$F_X^{-1}(p) = c_p = \inf\{x \in \mathbb{R} : F_X(x) \ge p\} = \begin{cases} 1, & 0 < p \le \frac{1}{6} \\ 2, & \frac{1}{6} < p \le \frac{1}{2} \\ 3, & \frac{1}{2} < p \le 1. \end{cases}$$

Remark 5. The end points of the intervals in the quantile function are the same as the p values of the CDF at the jumps. Furthermore, the $<$ inequality is always on the left of the x and the $\le$ inequality is always to the right of the x. This implies the quantile function is left continuous. Furthermore, the value of the quantile function on each interval is equal to the value of the quantile function at the right endpoint .

Remark 6. To find individual points of the quantile at p, we find the smallest point where the graph $F_X(x)$ lies on or above the horizontal line p. This is demonstrated for $p \in (0, 1/6]$ (left), $p \in (1/6, 1/2]$ (middle) and $p \in (1/2, 1]$ (right).



1.4.2 Derivations and Proofs

Problem 1.7. Prove Theorem 1.

Solution 1.7. This follows directly from the fact that X is discrete, so $\mathbb{P}(Y = y) = \mathbb{P}(g(X) = y)$ can be found by summing up all the probabilities of the values of x such that $g(x) = y$,

$$f_Y(y) = \mathbb{P}(Y = y) = \mathbb{P}(g(X) = y) = \sum_{x:g(x)=y} \mathbb{P}(X = x) = \sum_{x\in g^{-1}(\{y\})=y} f_X(x).$$

Problem 1.8. Prove Theorem 2.

Solution 1.8.

Strictly Increasing: We first consider the case that g is strictly increasing. If g is strictly increasing, then

$$F_Y(y) = \mathbb{P}(g(X) \le y) = \mathbb{P}(X \le g^{-1}(y)) = \int_{-\infty}^{g^{-1}(y)} f_X(t)\,dt.$$

Therefore, we can use the fundamental theorem of calculus and the chain rule to see that for points in the interior of the support of Y,

$$f_Y(y) = \frac{d}{dy}\int_{-\infty}^{g^{-1}(y)} f_X(t)\,dt = f_X(g^{-1}(y))\frac{d}{dy}g^{-1}(y) = \frac{1}{g'(g^{-1}(y))}f_X(g^{-1}(y)).$$

Since g is strictly increasing, $g' > 0$, so

$$f_Y(y) = \frac{1}{|g'(g^{-1}(y))|}f_X(g^{-1}(y)), \quad y \in \operatorname{supp}(Y).$$

Strictly Decreasing: We now consider the case that g is strictly decreasing. If g is strictly decreasing, then

$$F_Y(y) = \mathbb{P}(g(X) \le y) = \mathbb{P}(X \ge g^{-1}(y)) = 1 - \mathbb{P}(X \le g^{-1}(y)) = 1 - 1 - \int_{-\infty}^{g^{-1}(y)} f_X(t)\,dt.$$

Therefore, we can use the fundamental theorem of calculus and the chain rule to see that for points in the interior of the support of Y,

$$f_Y(y) = \frac{d}{dy}\left(1 - \int_{-\infty}^{g^{-1}(y)} f_X(t)\,dt\right) = -f_X(g^{-1}(y))\frac{d}{dy}g^{-1}(y) = -\frac{1}{g'(g^{-1}(y))}f_X(g^{-1}(y)).$$

Since g is strictly increasing, $g' < 0$, so

$$f_Y(y) = \frac{1}{|g'(g^{-1}(y))|}f_X(g^{-1}(y)), \quad y \in \operatorname{supp}(Y).$$

Problem 1.9. ($\star$) Prove the following properties for the quantile function

1. For all $x \in \mathbb{R}$, $F_X^{-1}(F_X(x)) \le x$
2. For all $p \in [0, 1]$, $F_X(F_X^{-1}(p)) \ge p$
3. $F_X^{-1}(p) \le x \Leftrightarrow p \le F_X(x)$
4. $F_X^{-1}(p)$ is non-decreasing and left-continuous (except for the endpoints $p = 0$ or $p = 1$)

Solution 1.9.

1. We have
$$F_X^{-1}(F_X(x)) = \inf_{t\in\mathbb{R}}\{F_X(t) \geq F_X(x)\} \leq x$$
since $x \in \{t \in \mathbb{R} : F_X(t) \geq F_X(x)\}$.

2. Since F_X is right continuous and increasing we have $\{F_X(x) \geq p\}$ is a closed set, so it attains its infimum. Therefore, $c_p \in \{F_X(x) \geq p\}$ so
$$F_X(F_X^{-1}(p)) = F_X(c_p) \geq p.$$

3. On one hand, $F_X^{-1}(p) \leq x$ implies that $x \in \{t : F_X(t) \geq p\}$ so $p \leq F_X(x)$. On the other hand, if $p \leq F_X(x)$ then $x \in \{t : F_X(t) \geq p\}$ so $F_X^{-1}(p) \leq x$ since $F_X^{-1}(p)$ is the infimum of all $\{t : F_X(t) \geq p\}$.

4. Suppose that $p_1 \leq p_2$. Then
$$F_X^{-1}(p_1) = \inf_{x\in\mathbb{R}}\{F_X(x) \geq p_1\} \leq \inf_{x\in\mathbb{R}}\{F_X(x) \geq p_2\} = F_X^{-1}(p_2)$$
since $\{F_X(x) \geq p_1\} \subseteq \{F_X(x) \geq p_2\}$, so F_X^{-1} is non-decreasing.

   To see left continuity, notice that monotone functions can only have jump discontinuities, so it suffices to show that $\sup_{q<p} F_X^{-1}(q) = F_X^{-1}(p)$. For each $q < p$ and $\epsilon > 0$, we have by definition of the supremum
$$\sup_{q<p} F_X^{-1}(q) + \epsilon \geq F_X^{-1}(q) \overset{(3)}{\Longrightarrow} F_X(\sup_{q<p} F_X^{-1}(q) + \epsilon) \geq q.$$
   So taking $\epsilon \to 0$ by right continuity of F_X implies that $F_X(\sup_{q<p} F_X^{-1}(q)) \geq q$ for all $q < p$ so $F_X(\sup_{q<p} F_X^{-1}(q)) \geq p$. Property 3 above implies that
$$\sup_{q<p} F_X^{-1}(q) \geq F_X^{-1}(p).$$
   This combined with monotonicity $\sup_{q<p} F_X^{-1}(q) \leq F_X^{-1}(p)$ implies that $\sup_{q<p} F_X^{-1}(q) = F_X^{-1}(p)$ as required.

2 Important Continuous Random Variables

2.1 Uniform Distribution: $\mathrm{U}(a, b)$

The uniform distribution models variables with equally likely outcomes on an interval. This is the continuous analogue of the discrete uniform distribution.

Definition 3. We say that X has a *continuous uniform distribution* on (a, b) if X has PDF

$$f_X(x) = \begin{cases} \frac{1}{b-a} & x \in (a,b), \\ 0 & \text{otherwise} \end{cases}$$

This is denoted by

$$X \sim \mathrm{U}(a, b).$$

Example 1. The following experiments can be modeled by a continuous uniform distribution:

Experiment	X	Distribution
Cutting a stick of length 1 at a random position	the location of the cut	$\mathrm{U}(0,1)$
Spinning a wheel	Location of spinner	$\mathrm{U}(0,2\pi)$

2.1.1 Properties

1. Sampling uniformly on the intervals (a, b), $[a, b)$, $(a, b]$ and $[a, b]$ are all equivalent.
2. **Mean and Variance:** If $X \sim \mathrm{U}(a, b)$ then

$$\mathbb{E}[X] = \frac{a+b}{2} \qquad \mathrm{Var}(X) = \frac{(b-a)^2}{12}$$

2.2 Exponential Distribution: $\mathrm{Exp}(\theta)$

The exponential distribution models the time between occurrences of a Poisson process with rate parameter $\lambda = \frac{1}{\theta}$ expressed in occurrences per time. The waiting time parameter (also called the mean parameter) θ is expressed in time per occurrence. This is the continuous time analogue of the geometric distribution.

Definition 4. We say that X has an exponential distribution with *mean* parameter θ if X has PDF

$$f_X(x) = \begin{cases} \frac{1}{\theta} e^{-\frac{x}{\theta}} & x > 0, \\ 0 & x \leq 0 \end{cases}$$

This is denoted by $X \sim \mathrm{Exp}(\theta)$.

Example 2. The following experiments can be modeled by a exponential distribution:

Experiment	X	Distribution
Busses arriving at rate 3 per hour	hours until a bus arrives	$\mathrm{Exp}(3^{-1})$
Busses arriving every 15 minutes	minutes until a bus arrives	$\mathrm{Exp}(15)$
Emails arrive at a rate of 20 per hour	hours until the next email	$\mathrm{Exp}(20^{-1})$

Remark 7. We can also parameterize the exponential random variable with the rate parameter. We say that X has an exponential distribution with rate parameter λ if X has PDF

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0, \\ 0 & x \leq 0 \end{cases}$$

This is (unfortunately also) denoted by $X \sim \mathrm{Exp}(\lambda)$. Unless otherwise stated, we will take the mean parametrization as the standard one in this course since that is the one given in your formula sheet.

2.2.1 Properties

1. **Parameters:** If a Poisson process has rate λ occurrences per time, then $\theta = \frac{1}{\lambda}$ is the time per occurrence.

2. **Mean and Variance:** If $X \sim \text{Exp}(\theta)$ then

$$\mathbb{E}[X] = \theta \qquad \text{Var}(X) = \theta^2$$

3. **Memoryless Property:** The exponential distribution forgets how long we have waited already.

Theorem 3 (***Memoryless property of Exp(θ)***)

If $X \sim \text{Exp}(\theta)$, *then*

$$\mathbb{P}(X > s + t | X > s) = \mathbb{P}(X > t).$$

In fact, the exponential distribution is the only continuous random variable with this property.

4. Computations with the exponential distribution can be expressed using the Gamma function

Definition 5 (Gamma function)**.** The integral

$$\Gamma(\alpha) = \int_0^\infty y^{\alpha-1} e^{-y} dy, \ \alpha > 0$$

is called the gamma function of α.

It satisfies the following nice properties

- $\Gamma(n) = (n-1)!$ for $n \in \mathbb{N}$
- $\Gamma(1/2) = \sqrt{\pi}$
- $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$ for $\alpha > 1$
- The Gamma function is a continuous function that interpolates the factorial function.

2.3 Example Problems

2.3.1 Applications

Problem 2.1. Suppose $X \sim \text{U}(0, 1)$, and that $Y = \frac{2}{X} - 1$. What is the support of Y?

Solution 2.1. Since $\text{supp}(X) = [0, 1]$, we have

$$0 \le X \le 1 \implies \infty > \frac{2}{X} \ge 2 \implies \infty > \frac{2}{X} - 1 \ge 1$$

so $\text{supp}(Y) = [1, \infty)$.

Problem 2.2. Suppose that the angle measured from the principal axis to the point of a spinner is uniformly distributed on $[0, 2\pi]$. You win the prize you want if the point lands in $[\frac{3\pi}{4}, \frac{3\pi}{2}]$. Given that the point will stop in the bottom half of the circle, what is the probability that you win the prize you want.

Solution 2.2. If X denotes the angle of the spinning wheel, then $X \sim U(0, 2\pi)$. Note that the bottom half is the circle $[\pi, 2\pi]$. Denote by A_1, A_2 the events $A_1 = \{X \in [\pi, 2\pi]\}$ and $A_2 = \{X \in [\frac{3\pi}{4}, \frac{3\pi}{2}]\}$. The desired probability is

$$\begin{aligned}
\mathbb{P}(A_2 \mid A_1) = \frac{\mathbb{P}(A_1 \cap A_2)}{\mathbb{P}(A_1)} &= \frac{\mathbb{P}(X \in [\pi, 2\pi] \cap [\frac{3\pi}{4}, \frac{3\pi}{2}])}{\mathbb{P}(X \in [\pi, 2\pi])} \\
&= \frac{\mathbb{P}(X \in [\pi, \frac{3\pi}{2}])}{\mathbb{P}(X \in [\pi, 2\pi])} \\
&= \frac{\frac{1}{2\pi} \cdot (\frac{3\pi}{2} - \pi)}{\frac{1}{2\pi} \cdot (2\pi - \pi)} \\
&= \frac{1/4}{1/2} = 1/2
\end{aligned}$$

Problem 2.3. Nupur decided to enjoy a relaxing Summer away from student housing, so she rented a place in Simcoe, Ontario. However, the busses there are far and few between. Suppose busses arrive according to a Poisson process with an average of 3 busses per hour.

1. Find the probability of waiting at least 15 minutes.
2. Find the probability of waiting at least another 15 minutes given that you have already been waiting for 6 minutes.

Solution 2.3.

Part 1: If X denotes the waiting time until the first bus in minutes, then the waiting time parameter is $\theta = 20$ min per bus so $X \sim \text{Exp}(20)$, and

$$\mathbb{P}(X \geq 15) = \int_{15}^{\infty} \frac{1}{20} e^{-\frac{1}{20}x} dx = e^{-\frac{3}{60} \cdot 15} = e^{-3/4}$$

Alternative Solution: If X denotes the waiting time until the first bus in hours, then the waiting time parameter is $\theta = \frac{1}{3}$ hours per bus (rate λ is 3 busses per hour) $X \sim \text{Exp}(3^{-1})$. Hence, the probability of waiting at least 15 minutes is

$$\mathbb{P}(X \geq 1/4) = \int_{1/4}^{\infty} 3e^{-3x} dx = e^{-3 \cdot \frac{1}{4}}.$$

Part 2: By the memoryless property, we have

$$\mathbb{P}(X \geq 15 + 6 \mid X \geq 6) = \mathbb{P}(X \geq 15) = e^{-3/4}$$

from above.

Problem 2.4. Exponential distribution is also very useful in reliability engineering. The lifetime of a seat belt motor on a 1994 Saturn GL is known to follow an exponential distribution with mean 14 years.

1. What is the standard deviation of the lifetime of a seat belt motor on a 1994 Saturn GL?
2. Compute the probability that the lifetime of the seat belt motor will last more than 6 years.
3. If a seat belt motor has lasted 14 years, what is the probability that it will last another 6 years?

Solution 2.4. In this problem we are given $\mathbb{E}[X] = 14$ and X has exponential distribution so $X \sim \text{Exp}(14)$.

1. We have $\text{Var}(X) = 14^2$, so the standard deviation is 14 years.
2. We have
$$\mathbb{P}(X \geq 6) = \int_6^\infty \frac{1}{14} e^{-\frac{1}{14}t}\, dt = e^{-\frac{6}{14}}.$$
3. By the memoryless property, we have
$$\mathbb{P}(X \geq 6 + 14 \mid X \geq 14) = \mathbb{P}(X \geq 6) = e^{-\frac{6}{14}}.$$

Problem 2.5. Suppose the waiting time X until Mukhtar's next bus arrives follows an exponential distribution with parameter $\theta = 1$. What's the waiting time w so that Mukhtar doesn't have to wait longer than w with probability 50%?

Solution 2.5. We know
$$F_X(x) = \mathbb{P}(X \leq x) = \int_0^x e^{-t}\, dt = 1 - e^{-x}$$
for $x \geq 0$ and 0 otherwise. This function is strictly increasing on $x > 0$, so we can use the classical inverse. We want w such that $\mathbb{P}(X \leq w) = 0.5$, or
$$F_X(w) = 0.5 \Leftrightarrow 1 - e^{-w} = 0.5 \Leftrightarrow w = \log(2) \approx 0.693.$$
So with probability 50% Mukhtar won't have to wait longer than $\log(2)$. In other words, $\log(2)$ is the median of the distribution of X.

Remark 8. We can repeat this computation for general $\theta \neq 1$. The median in this case will be given by $\theta \log(2)$. Since $\log(2) < 1$ this implies that the median of the exponential is always below the mean.

Problem 2.6. Uranium 238 emits particles measured by a Geiger counter at a rate of 50 per second. Assume that the number of particles measured by a Geiger counter follows a Poisson process. Let X denote the amount of time in seconds between when the first and second particles are measured. Find $\mathbb{E}[X]$

Solution 2.6. By the homogeneous property and independence, the time between the first and second particle is equal in distribution to the time between the first particle. Therefore, the waiting time parameter θ is $\frac{1}{50}$ seconds per occurrence, so $X \sim \text{Exp}(50^{-1})$. Therefore, $\mathbb{E}[X] = \theta = \frac{1}{50}$.

Careful Solution: We carefully explain how the homogeneous and independence property are used in this problem. Let X be the time of the first particle, Y be the time from the first to the second particle, and $N(t)$ be the number of particles by time t. Clearly, $Z = X + Y$ is the time of the second particle.

We know $X \sim \text{Exp}(50^{-1})$ and $N(t) \sim \text{Poi}(50t)$. Using the same logic as in the derivation of the Poisson process (see Problem (2.9)),
$$F_Z(t) = \mathbb{P}(X + Y \leq t) = 1 - \mathbb{P}(X + Y > t) = 1 - P(N(t) \leq 1) = 1 - (e^{-50t} + e^{-50t}(50t))$$
So $Z = X + Y$ has density
$$f_Z(t) = \frac{d}{dt}\mathbb{P}(X + Y \leq t) = -\frac{d}{dt}\left[e^{-50t}(1 + 50t)\right] = t50^2 e^{-50t}$$

Therefore, by an integration by parts

$$\mathbb{E}[Z] = \int_{-\infty}^{\infty} z \cdot f_Z(z) dz = 50 \int_0^{\infty} z^2 \cdot 50e^{-50z} dz = 50\frac{2}{50^2} = \frac{2}{50}$$

and the linearity of expectation implies that

$$\mathbb{E}[Y] = \mathbb{E}[Z - X] = \mathbb{E}[Z] - \mathbb{E}[X] = \frac{2}{50} - \frac{1}{50} = \frac{1}{50}.$$

Problem 2.7. A continuous random variable X is said to have a Gamma distribution with parameters $\alpha > 0$ and $\beta > 0$ if it has PDF

$$f_X(x) = \begin{cases} \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-\frac{x}{\beta}} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

Use the properties of the Gamma function, namely $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$ for $\alpha > 1$, to obtain $\mathbb{E}[X]$ and $\mathrm{Var}(x)$.

Solution 2.7. We can solve this problem without even knowing what the $\Gamma(\alpha)$ function is. We use a trick that we have used many times, which involves rewriting the expression as an integral of the PDF which sums to 1.

Expected Value: By definition,

$$\begin{aligned} \mathbb{E}[X] = \int_0^{\infty} x \cdot \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-\frac{x}{\beta}} \, dx &= \int_0^{\infty} \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha} e^{-\frac{x}{\beta}} \, dx \\ &= \frac{\beta\Gamma(\alpha+1)}{\Gamma(\alpha)} \underbrace{\int_0^{\infty} \frac{1}{\Gamma(\alpha+1)\beta^{\alpha+1}} x^{\alpha} e^{-\frac{x}{\beta}} \, dx}_{=1 \text{ integral of the PDF of } \Gamma(\alpha+1,\beta) \text{ r.v.}} \end{aligned}$$

where we multiplied and divided by $\beta\Gamma(\alpha + 1)$ to match the normalization terms. Since using the recursive property of the Γ function, we see that

$$\mathbb{E}[X] = \frac{\beta\Gamma(\alpha+1)}{\Gamma(\alpha)} = \frac{\beta\alpha\Gamma(\alpha)}{\Gamma(\alpha)} = \beta\alpha.$$

Variance: A similar computation as above implies that

$$\begin{aligned} \mathbb{E}[X^2] = \int_0^{\infty} x^2 \cdot \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-\frac{x}{\beta}} \, dx &= \int_0^{\infty} x^2 \cdot \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha+1} e^{-\frac{x}{\beta}} \, dx \\ &= \frac{\beta^2\Gamma(\alpha+2)}{\Gamma(\alpha)} \underbrace{\int_0^{\infty} \frac{1}{\Gamma(\alpha+2)\beta^{\alpha+2}} x^{\alpha+1} e^{-\frac{x}{\beta}} \, dx}_{=1 \text{ integral of the PDF of } \Gamma(\alpha+2,\beta) \text{ r.v.}} \\ &= \frac{\beta^2\Gamma(\alpha+2)}{\Gamma(\alpha)}. \end{aligned}$$

Next, using the recursive properties of the Γ function,

$$\mathbb{E}[X^2] = \frac{\beta^2\Gamma(\alpha+2)}{\Gamma(\alpha)} = \frac{\beta^2(\alpha+1)\Gamma(\alpha+1)}{\Gamma(\alpha)} = \frac{\beta^2(\alpha+1)\alpha\Gamma(\alpha)}{\Gamma(\alpha)} = \beta^2(\alpha+1)\alpha.$$

Therefore,

$$\mathrm{Var}(x) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \beta^2(\alpha+1)\alpha - \beta^2\alpha^2 = \beta^2\alpha.$$

Remark 9. Notice that then $\alpha = 1$, then the PDF is simply the exponential distribution. In fact, the relation is even deeper and it can be shown that the sum of n independent $\mathrm{Exp}(\beta)$ distributions has a $\Gamma(n, \beta)$ distribution.

2.3.2 Derivations and Proofs

Problem 2.8. Compute the mean and variance of $X \sim \mathrm{U}(a, b)$.

Solution 2.8. We can directly compute

$$\begin{aligned}\mathbb{E}(X) &= \int_{-\infty}^{\infty} x f(x)\, dx = \int_a^b x \cdot \frac{1}{b-a}\, dx \\ &= \frac{1}{2}\frac{b^2 - a^2}{b-a} = \frac{(b-a)(b+a)}{2(b-a)} = \frac{a+b}{2}.\end{aligned}$$

To compute the variance, we have

$$\begin{aligned}\mathbb{E}(X^2) &= \int_{-\infty}^{\infty} x^2 f(x)\, dx = \int_a^b x^2 \frac{1}{b-a}\, dx \\ &= \frac{b^3 - a^3}{3(b-a)} = \frac{(b-a)(b^2+ab+a^2)}{3(b-a)} = \frac{b^2+ab+a^2}{3}.\end{aligned}$$

so after some algebra,

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - \mathbb{E}(X)^2 = \frac{b^2+ab+a^2}{3} - \frac{(a+b)^2}{4} = \frac{(b-a)^2}{12}.$$

Problem 2.9. Let $N(t)$ be a Poisson process with rate λ occurrences per time interval. If X is the length of time until the first occurrence, show that $X \sim \mathrm{Exp}(\lambda^{-1})$.

Solution 2.9. Since $\mathbb{P}(X \leq 0) = 0$, it remains to consider $x > 0$. We start by computing the CDF of X, which is the length of time until first event occurs

$$\begin{aligned}F_X(t) &= \mathbb{P}(X \leq t) \\ &= \mathbb{P}(\text{time to 1st occurrence} \leq t) \\ &= 1 - \mathbb{P}(\text{time to first occurrence} > t) \\ &= 1 - \mathbb{P}(\text{no occurrence between } (0,t))\end{aligned}$$

We know how to model the number of event occurrences between time $(0, t)$ since it is $\mathrm{Poi}(\lambda t)$. Let $N(t) \sim \mathrm{Poi}(\lambda t)$. Then

$$\begin{aligned}F_X(t) &= 1 - \mathbb{P}(\text{no occurrence between } (0,t)) \\ &= 1 - \mathbb{P}(N(t) = 0) \\ &= 1 - \frac{e^{-\lambda t}(\lambda t)^0}{0!} \\ &= 1 - e^{-\lambda t}.\end{aligned}$$

So we have $F_X(t) = 1 - \exp(-\lambda t)$ for $t > 0$ and $F_X(t) = 0$ otherwise. We can take the derivative with respect to t for $t > 0$, to obtain the PDF

$$f_X(t) = \frac{d}{dt} F_X(t) = \lambda \exp(-\lambda t)$$

and $f_X(t) = 0$ for $t \leq 0$. We recognize this PDF as the one corresponding to $\mathrm{Exp}(\lambda^{-1})$ (with waiting time parameter $\theta = \frac{1}{\lambda}$).

Remark 10. The (actual) Poisson process $N(t)$ counts the number of occurrences up to time t. While X is the waiting time until the first occurrence. By independent increments and homogeneous property, the waiting time between occurrences has the same distribution as the waiting time until the first occurrence. This means that the exponential distribution $\mathrm{Exp}(\lambda^{-1})$ models the waiting time between each event occurrence in a Poisson distribution with rate λ.

Problem 2.10. Let $n \geq 1$. Show that

$$\Gamma(n) = \int_0^\infty y^{n-1} e^{-y} dy = (n-1)!.$$

Solution 2.10. This follows from repeated integration by parts. Let $n \geq 1$, we have

$$\Gamma(n) = \int_0^\infty y^{n-1} e^{-y} dy.$$

Integrating by parts we see that

$\pm$	D	I
$+$	y^{n-1}	e^{-y}
$-$	$(n-1)y^{n-2}$	$-e^{-y}$

so

$$\Gamma(n) = \int_0^\infty y^{n-1} e^{-y} dy = -y^{n-1}e^{-y}\Big|_0^\infty + (n-1)\int_0^\infty y^{n-2}e^{-y} dy = (n-1)\Gamma(n-1).$$

Repeating this inductively implies that

$$\Gamma(n) = (n-1)\Gamma(n-1) = (n-1)(n-2)\Gamma(n-2) = \cdots = (n-1)(n-2)\cdots 3 \cdot 2 \cdot \Gamma(1) = n!$$

since $\Gamma(1) = 1$ (using the fact that the integral of the PDF of $X \sim \mathrm{Exp}(1)$ is 1).

Remark 11. Notice that the integration by parts computation holds for $\alpha > 1$ which are not necessarily integer valued to conclude that $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$. The problem is that the inductive step might not lead to a nice number since $\Gamma(x)$ for $x < 1$ doesn't necessarily simplify.

Problem 2.11. Compute the mean and variance of $X \sim \mathrm{Exp}(\theta)$.

Solution 2.11. We use the change of variable $x = y\theta$ with $dx = \theta dy$

$$\begin{aligned}\mathbb{E}[X] &= \int_0^\infty x \cdot \frac{1}{\theta} e^{-\frac{x}{\theta}} dx \overset{(x=y\theta)}{=} \int_0^\infty y e^{-y} \theta dy \\ &= \theta \underbrace{\int_0^\infty y e^{-y} dy}_{=\Gamma(2)} = \theta\Gamma(2) = \theta \cdot (1!) = \theta.\end{aligned}$$

To compute the variance, we have

$$\begin{aligned}\mathbb{E}[X^2] &= \int_0^\infty x^2 \cdot \frac{1}{\theta} e^{-\frac{x}{\theta}} dx \overset{(x=y\theta)}{=} \int_0^\infty \theta y^2 e^{-y} \theta dy \\ &= \theta^2 \underbrace{\int_0^\infty y^{3-1} e^{-y} dy}_{=\Gamma(3)} = \theta^2\Gamma(3) = \theta \cdot (2!) = 2\theta^2.\end{aligned}$$

so

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

Alternative Solution: We can also integrate by parts directly without using the Gamma function. Starting with

$$\mathbb{E}[X] = \int_0^\infty x \cdot \frac{1}{\theta} e^{-\frac{x}{\theta}}\, dx$$

we have by integrating by parts

$\pm$	D	I
$+$	$\frac{x}{\theta}$	$e^{-\frac{x}{\theta}}$
$-$	$\frac{1}{\theta}$	$-\theta e^{-\frac{x}{\theta}}$
$+$	0	$\theta^2 e^{-\frac{x}{\theta}}$

so

$$\mathbb{E}[X] = \int_0^\infty x \cdot \frac{1}{\theta} e^{-\frac{x}{\theta}}\, dx = \frac{x}{\theta} \cdot (-\theta e^{-\frac{x}{\theta}}) - \frac{1}{\theta}\theta^2 e^{-\frac{x}{\theta}} \Big|_0^\infty = \theta.$$

Next, to compute

$$\mathbb{E}[X^2] = \int_0^\infty x^2 \cdot \frac{1}{\theta} e^{-\frac{x}{\theta}}\, dx$$

we have by integrating by parts

$\pm$	D	I
$+$	$\frac{x^2}{\theta}$	$e^{-\frac{x}{\theta}}$
$-$	$\frac{2x}{\theta}$	$-\theta e^{-\frac{x}{\theta}}$
$+$	$\frac{2}{\theta}$	$\theta^2 e^{-\frac{x}{\theta}}$
$-$	0	$-\theta^3 e^{-\frac{x}{\theta}}$

so

$$\mathbb{E}[X^2] = \int_0^\infty x^2 \cdot \frac{1}{\theta} e^{-\frac{x}{\theta}}\, dx = \frac{x^2}{\theta} \cdot (-\theta e^{-\frac{x}{\theta}}) - \frac{2x}{\theta} \cdot \theta^2 e^{-\frac{x}{\theta}} + \frac{2}{\theta} \cdot (-\theta^3 e^{-\frac{x}{\theta}}) \Big|_0^\infty = 2\theta^2.$$

Problem 2.12. Prove the memoryless property: Theorem 3.

Solution 2.12. Recall the CDF of $X \sim \mathrm{Exp}(\theta)$ is

$$F_X(x) = \mathbb{P}(X \le x) = \int_{-\infty}^x f_X(t) dt = \int_0^x \theta^{-1} e^{-t/\theta}\, dt = 1 - e^{-x/\theta}, \quad x > 0.$$

Therefore,

$$\mathbb{P}(X > x) = 1 - \mathbb{P}(X \le x) = 1 - F_X(x) = e^{-x/\theta}.$$

Hence,

$$\begin{aligned}\mathbb{P}(X > s + t \mid X > s) &= \frac{\mathbb{P}(X > s+t \text{ and } X > s)}{\mathbb{P}(X > s)} \\ &= \frac{\mathbb{P}(X > s+t)}{\mathbb{P}(X > s)} = \frac{e^{-(s+t)/\theta}}{e^{-s/\theta}} \\ &= e^{-t/\theta} = \mathbb{P}(X > t)\end{aligned}$$

as desired.

Problem 2.13. ($\star$) Prove that $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

Solution 2.13. We have

$$\Gamma\Big(\frac{1}{2}\Big) = \int_0^\infty y^{\frac{1}{2}-1} e^{-y} dy = \int_0^\infty y^{-\frac{1}{2}} e^{-y} dy$$

We can do the change of variables $y = \frac{x^2}{2}$, $dy = x\,dx$ so

$$\Gamma\Big(\frac{1}{2}\Big) = \sqrt{2}\int_0^\infty e^{-\frac{x^2}{2}}\,dx.$$

By Fubini's theorem and a change of variables into polar coordinates

$$\begin{aligned}\Gamma\Big(\frac{1}{2}\Big)^2 = 2\bigg(\int_0^\infty e^{-\frac{x^2}{2}}\,dx\bigg)^2 = 2\int_0^\infty e^{-\frac{x^2}{2}}\,dx \int_0^\infty e^{-\frac{y^2}{2}}\,dy &= 2\int_0^\infty\int_0^\infty e^{-\frac{x^2+y^2}{2}}\,dxdy \\ &= 2\int_0^{\frac{\pi}{2}} e^{-\frac{r^2}{2}} r\,drd\theta \\ &= 2\int_0^{\frac{\pi}{2}} d\theta = \pi.\end{aligned}$$

So $\Gamma\Big(\frac{1}{2}\Big) = \sqrt{\pi}$.

Remark 12. A modification of this argument can be used to compute the normalization constant of a standard Gaussian random variable we will see next week.