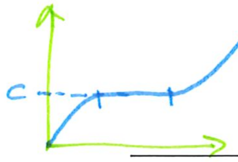
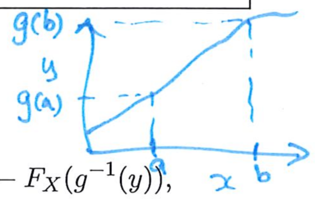


$Y = g(X)$ $\qquad$ $\mathbb{P}(Y = c) > 0$

$$\boxed{\frac{d}{dy}F_Y(y) = \frac{d}{dy} F_X(g^{-1}(y)) = \frac{1}{g'(y)} f_X(g^{-1}(y))}$$

If X has support $[a, b]$, then the support of Y is $\boxed{[g(a), g(b)]}$.



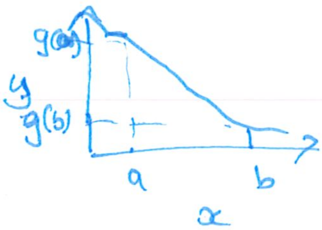
On the other hand, if g is decreasing, then

$$F_Y(y) = \mathbb{P}(Y \leqslant y) = \mathbb{P}(X \geqslant g^{-1}(y)) = 1 - \lim_{z \uparrow y} F_X(g^{-1}(z)) = 1 - F_X(g^{-1}(y)),$$

because g^{-1} is decreasing

The pdf of Y is then given by

$$\frac{d}{dy}F_Y(y) = \frac{d}{dy}\left(1 - F_X(g^{-1}(y))\right) = \frac{1}{|g'(y)|} f_X(g^{-1}(y)).$$



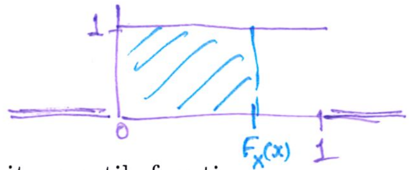
If X has support $[a, b]$, then the support of Y is $\boxed{[g(b), g(a)]}$.

An important type of monotone transformation is given by the *inverse cdf* or *quantile function*.

Defintion. Let X be a continuous random variable. The function $q_X : \mathbb{R} \to \mathbb{R}$ such that

$$F_X(q_X(x)) = x,$$

is called the **quantile function** of X.



Note that the quantile function is an increasing function.

Suppose F_X is the cdf of a continuous random variable X and q_X is its quantile function. If $U \sim \mathsf{Uniform}(0, 1)$, then the cdf of $q_X(U)$ is

$$\boxed{\begin{aligned}\mathbb{P}(q_X(U) \leqslant x) &= \mathbb{P}(F_X(q_X(U)) \leqslant F_X(x)) \\ &= \mathbb{P}(U \leqslant F_X(x)) = F_X(x)\end{aligned}}$$

Example. For $X \sim \mathsf{Exp}(\lambda)$ (consider Figure 5.4) we have:

$$\begin{aligned} x &= F_X(q) = 1 - e^{-\lambda q} \\ 1 - x &= e^{-\lambda q} \\ \log(1-x) &= -\lambda q \\ \frac{-\log(1-x)}{\lambda} &= q \\ \Rightarrow \quad q_X(x) &= -\frac{\ln(1-x)}{\lambda} \end{aligned}$$

when $x \geqslant 0$. Note that if $U \sim \mathsf{Uniform}(0, 1)$, then $V = 1 - U$ has a $\boxed{\text{Uniform}(0,1)}$ distribution.

As a result we can define a MATLAB function to generate samples from the Exponential distribution as follows.

$v \in (0,1)$

$$\begin{aligned} F_V(v) &= \mathbb{P}(V \leqslant v) \\ &= \mathbb{P}(1 - U \leqslant v) \\ &= \mathbb{P}(1 - v \leqslant U) \\ &= 1 - \mathbb{P}(U < 1 - v) = 1 - (1 - v) = v \end{aligned}$$

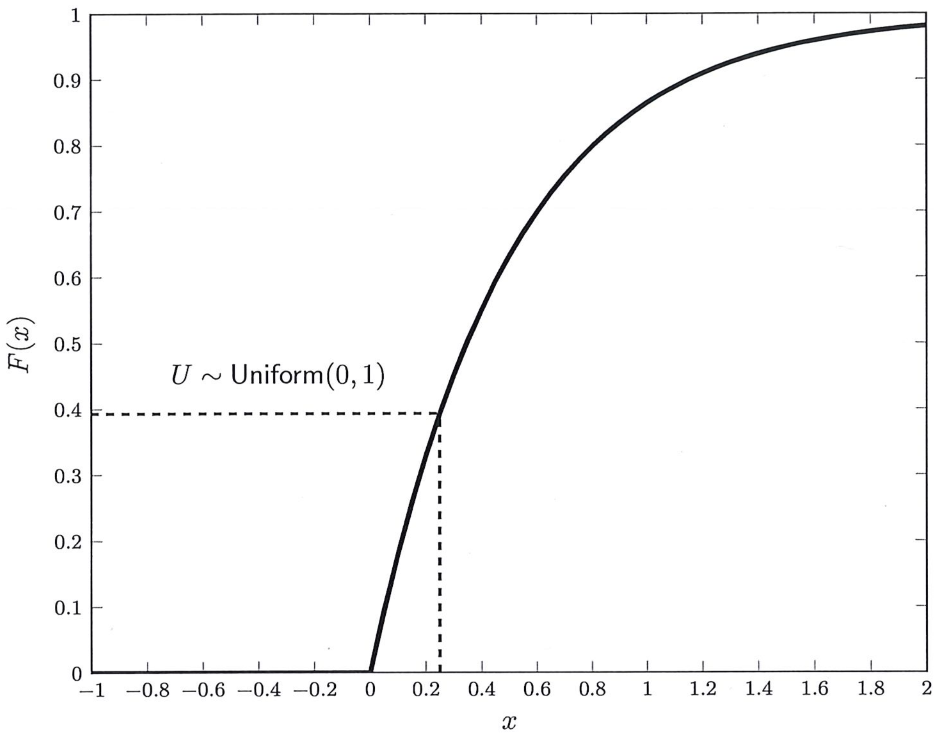


Figure 5.4: The cdf of $X \sim \mathsf{Exp}(2)$.

```
function output = Exponential(lambda)
    output = -log(rand)/lambda;
end
```

Upon saving this as 'Exponential.m' to our working directory we can then use this function as follows:

```
>> Exponential(2)
ans =
    0.0453
>> Exponential(2)
ans =
    0.2291
>> Exponential(2)
ans =
    1.1637
```

Be careful, as the built in MATLAB function **exprnd** generates samples from the $\mathsf{Exp}(\lambda^{-1})$ distribution.

Non-monotone transformations

If the function g is not monotone, then we can still make progress by considering separately those intervals over which it is monotone. The general procedure to follow is given below:

$Y = g(X)$ g is non monotone.

1. Determine the support of the Y.

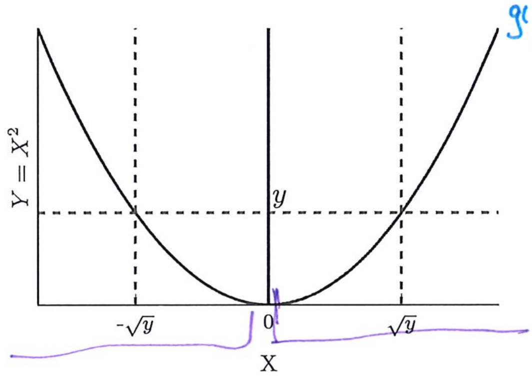
2. Determine the event in terms of the random variable X that maps to the event $\{Y \leqslant y\}$. Typically this will be in the form of a union of disjoint events of the form $\{a \leqslant X \leqslant b\}$.

3. Find the probability $F_Y(y)$ of the event $\{Y \leqslant y\}$ in terms of F_X, the cumulative distribution function of X.

4. Differentiate the result to find the probabilty density function of Y.

Example. Suppose $X \sim \mathsf{Normal}(0, 1)$ and $Y = X^2$. Find the pdf of Y.

here g is decreasing



here g is increasing

Step 1: The function $g(x) = x^2$ maps $\mathbb{R}$ to $[0, \infty)$. So the support of Y is $[0\infty)$.

Step 2: From the figure it is clear that $\{Y \leqslant y\}$, where $y \geqslant 0$, corresponds to $\{-\sqrt{y} \leqslant X \leqslant \sqrt{y}\}$.

Steps 3 and 4:

$y \geqslant 0$

$$\mathbb{P}(Y \leqslant y) = \mathbb{P}(-\sqrt{y} \leqslant X \leqslant \sqrt{y})$$
$$= F_X(\sqrt{y}) - F_X(-\sqrt{y})$$

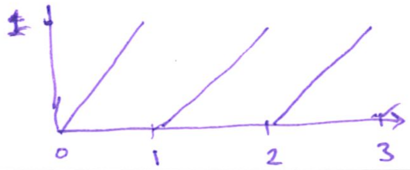
so that $f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy}\left(F_X(\sqrt{y}) - F_X(-\sqrt{y})\right)$

(chain rule)

$$f_Y(y) = \frac{1}{2\sqrt{y}} f_X(\sqrt{y}) + \frac{1}{2\sqrt{y}} f_X(-\sqrt{y})$$
$$= \frac{1}{\sqrt{2\pi y}} e^{-y/2}, \quad y \geqslant 0$$

This distribution is called the χ_1^2-distribution.

$$\frac{1}{2\sqrt{y}} \frac{1}{\sqrt{2\pi}} e^{-(\sqrt{y})^2/2} + \frac{1}{2\sqrt{y}} \frac{1}{\sqrt{2\pi}} e^{-(-\sqrt{y})^2/2}$$



Example. Suppose $X \sim \mathsf{Exp}(\lambda)$ and $Y = X - \lfloor X \rfloor$. Find the pdf of Y. (*Note that* Y *is the fractional part of* X.)

Let $y \in (0,1)$. The event $\{Y \leqslant y\}$ can be written in terms of the random variable X as

$$\bigcup_{n=1}^{\infty} \{n-1 \leqslant X \leqslant n-1+y\}$$

$$\{0 \leqslant X \leqslant y\} \cup \{1 \leqslant X \leqslant 1+y\} \cup \{2 \leqslant X \leqslant 2+y\}$$

$$\begin{aligned}
\mathbb{P}(Y \leqslant y) &= \mathbb{P}\left(\bigcup_{n=1}^{\infty} \{n-1 \leqslant X \leqslant n-1+y\}\right) \\
&= \sum_{n=1}^{\infty} \mathbb{P}(\{n-1 \leqslant X \leqslant n-1+y\}) \\
&= \sum_{n=1}^{\infty} \left[F_X(n-1+y) - F_X(n-1)\right]
\end{aligned}$$

if A and B are disjoint $P(A \cup B) = P(A) + P(B)$

so that

$$\begin{aligned}
f_Y(y) &= \frac{d}{dy} \sum_{n=1}^{\infty} \left[F_X(n-1+y) - F_X(n-1)\right] \\
&= \sum_{n=1}^{\infty} f_X(n-1+y) \\
&= \sum_{n=1}^{\infty} \lambda e^{-\lambda(n-1+y)} = \lambda e^{-\lambda y} \sum_{n=1}^{\infty} e^{-\lambda(n-1)} = \lambda e^{-\lambda y}(1-e^{-\lambda})^{-1}
\end{aligned}$$

Note. We have previously defined the expected value of a continuous random variable X and the expected value of a function g of X. Now that we have studied the effect of transformations on the distribution of a random variable we can see that these two definitions are consistent. That is, given a continuous random variable X and continouos function g, if we define the random variable $Y := g(X)$, then $\mathbb{E}Y = \mathbb{E}[g(X)]$.

Multiple continuous random variables

As was the case with discrete random variables, we will often have need to work with multiple random variables at once. Recall that the **joint distribution** of the random variables $X_1, \ldots, X_n$, defined in the same random experiment, can be specified through the **joint cumulative distribution function** F defined by

$$F(x_1, \ldots, x_n) = \mathbb{P}(X_1 \leqslant x_1, \ldots, X_n \leqslant x_n).$$

This completely specifies the probability distribution of the vector $\boldsymbol{X} := (X_1, \ldots, X_n)$.

We will now just work with a pair of continuous random variables (X, Y) having joint cdf $F_{X,Y}$. The extension to more than two random variables is straightforward.

Lets first recall some basic notions that we saw previously in connection with the joint distribution of multiple discrete random variables.

It is clear from the law of total probability that

$\mathbb{P}(X \leqslant x) = \mathbb{P}(X \leqslant x, Y < \infty)$

$$F_X(x) = \mathbb{P}(X \leqslant x) = \lim_{y \to \infty} F_{X,Y}(x, y)$$

and similarly for F_Y. We refer to F_X and F_Y as **marginal cumulative distribution functions.** From the joint cumulative distribution $F_{X,Y}$ we can determine if X and Y are independent: X and Y are said to be **independent** if

$$F_{X,Y}(x, y) = F_X(x) F_Y(y),$$

Recall X and Y are independent iff all pairs of events $\{X \leqslant x\}$ and $\{Y \leqslant y\}$ are independent.

for all $(x, y) \in \mathbb{R}^2$. Instead of using the cdf to describe the distribution of a single continuous random variable, we usually used its probability density function. Similarly, for multiple continuous random variables we usually use the joint probability density function.

Definition. If there exists a function $f_{X,Y}(x, y)$ such that for all $(x, y) \in \mathbb{R}^2$

$$F_{X,Y}(x, y) = \int_{-\infty}^{y} \int_{-\infty}^{x} f_{X,Y}(u, v)\, \mathrm{d}u\, \mathrm{d}v,$$

we call $f_{X,Y}$ the **joint probability density function** of (X, Y).

Note that in the above double integral we integrate the variable u first, treating v as constant, and then integrate the variable v. In this setting, the order in which we perform this integration is not important since it can be shown that

$$\int_{-\infty}^{y} \int_{-\infty}^{x} f_{X,Y}(u, v)\, \mathrm{d}u\, \mathrm{d}v = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(u, v)\, \mathrm{d}v\, \mathrm{d}u.$$

The joint pdf is not prescribed uniquely by this definition, but, if both of the *partial derivatives* of $F_{X,Y}$ exist at the point (x, y), then

$$f_{X,Y}(x, y) = \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x, y).$$

The symbol $\frac{\partial^2}{\partial x \partial y}$ means to differentiate $F_{X,Y}(x, y)$ first with respect to y, treating x as constant and then differentiate with respect to x, treating y as constant. The order in which we perform this differentiation is not imporant since it can be shown that

$$\frac{\partial^2}{\partial x \partial y} F_{X,Y}(x, y) = \frac{\partial^2}{\partial y \partial x} F_{X,Y}(x, y).$$

The joint pdf completely specifies the distribution of (X, Y), as does the joint cdf.

Basic properties of $f_{X,Y}$:

- $f_{X,Y}(x, y) \geqslant 0$ for all $(x, y) \in \mathbb{R}^2$;
- $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y)\, \mathrm{d}x \mathrm{d}y = 1.$ $\Longleftrightarrow \mathbb{P}(\Omega) = 1$
- $\mathbb{P}(a \leqslant X \leqslant b, c \leqslant Y \leqslant d) = \int_a^b \int_c^d f_{X,Y}(x, y)\, \mathrm{d}y\, \mathrm{d}x$, where a or c can be $-\infty$ and b or d can be ∞, and any of the inequalities can be replaced by strict ones.

Probabilities now given by volumes under the joint pdf.