1 Sampling Realizations of Random Variables

In practice, we may want to sample realizations of random variables.

Example 1. If $X \sim \text{Exp}(1)$, then we can estimate $\mathbb{P}(X > 2)$ by sampling $x_1, x_2, \ldots, x_n$ from a Exp(1) distribution, then approximate using the relative frequency of samples larger than 2,

$$\mathbb{P}(X > 2) \approx \frac{|\{x_1, \ldots, x_n : x_i > 2\}|}{n}.$$

It is "easy" to sample from the continuous uniform distribution $\mathrm{U}(0,1)$ on a computer. These uniform random variables can be used to generate samples from any distribution.

Theorem 1 (*Inverse Transform Sampling I*)

If F_X is a continuous and strictly increasing CDF of some random variable X and if $U \sim \mathrm{U}(0,1)$, then the random variable $Y = F_X^{-1}(U)$ has CDF F_X.

By using the more general quantile function

$$F_X^{-1}(y) = \inf\{x \in \mathbb{R} : F_X(x) \geq y\}$$

one can show the following generalization:

Theorem 2 (*Inverse Transform Sampling II*)

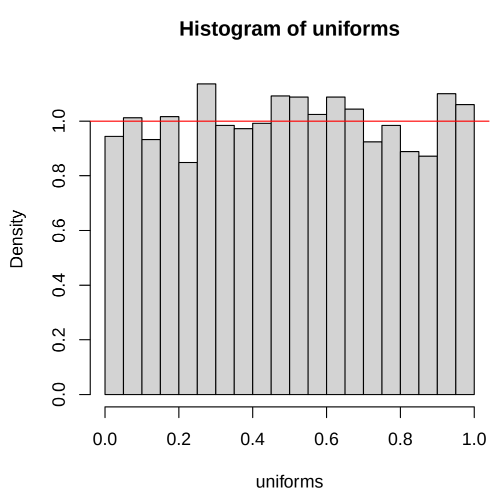
Let F_X be any cumulative distribution function of some random variable X and $U \sim U(0,1)$. Then the random variable $Y = F_X^{-1}(U)$ has CDF F_X.

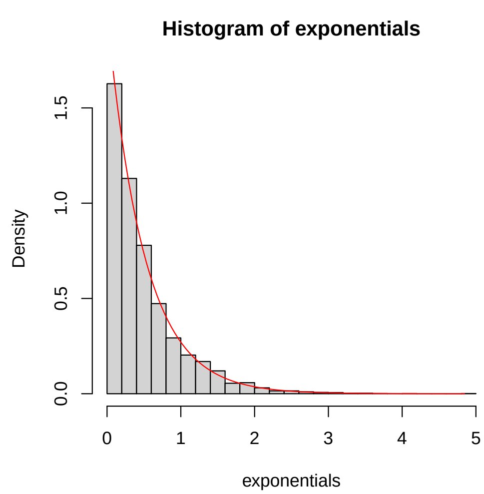
Remark 1. This is a generalization of a simple concept. For instance, if we want to generate a flip of a coin (a Ber(0.5) random variable), then we can sample a number uniformly u from [0.1] and define $x = X(u) = 0$ if $u_1 \in [0, 0.5]$ and $x = X(u) = 1$ if $u_1 \in [0.5, 1]$.

1.1 Sampling Algorithm

1. No matter what CDF F_X (discrete or continuous), we can sample observations as follows:
   (a) Sample $u \sim U(0,1)$ (eg via `runif()`)
   (b) Return $x = F_X^{-1}(u)$.
2. Repeating this n times independently gives n realizations of X.

Example 2. We sample `uniforms <- runif(5000)` and then `exponentials <- -log(1-uniforms)/2`.





1.2 Example Problems

1.2.1 Applications

Problem 1.1. Let $U \sim \mathrm{U}(0,1)$. We want to sample from the $\mathrm{Exp}(2^{-1})$ distribution with density

$$f_X(x) = 2e^{-2x}, \quad x > 0$$

and 0 otherwise. Write Y as a function of U such that Y is equal in distribution to X.

Solution 1.1. The CDF on the support of X

$$F_X(x) = \int_0^x 2e^{-2t} dt = 1 - e^{-2x},$$

which is strictly increasing for on its support $x \geq 0$. Solving for $F_X(y) = x$ to recover the inverse gives $y = F_X^{-1}(x) = -\frac{1}{2}\log(1-x)$, so

$$F_X^{-1}(x) = -\frac{1}{2}\log(1-x)$$

for $x \in (0,1)$. Therefore, by Theorem 1 (or Theorem 2)

$$Y = -\frac{1}{2}\log(1-U)$$

has the same distribution $Y \sim \mathrm{Exp}(2^{-1})$.

Problem 1.2. Suppose that we wish to generate a random observation, x, from a distribution with PDF given by

$$f_X(x) = \frac{1}{8\sqrt{x}}, \quad 0 < x < 16$$

and 0 otherwise. We generate an observation, u, from a continuous $\mathrm{U}(0,1)$ distribution (using software) and get 0.1348. Determine the value $x = x(u)$, that this value u will produce.

Solution 1.2. We first compute the CDF on the support of X

$$F_X(x) = \int_0^x \frac{1}{8\sqrt{t}}\, dt = \frac{1}{4}\sqrt{x}, \quad 0 < x < 16.$$

which is strictly increasing on its support $0 < x < 16$. Solving for $F_X(y) = x$ to recover the inverse gives $y = F^{-1}(x) = (4x)^2$, so

$$F_X^{-1}(x) = (4x)^2$$

for $x \in (0,1)$. By the sampling algorithm, if $u = 0.1348$ the corresponding observation of x is

$$x = F_X^{-1}(u) = (4 \cdot 0.1348)^2 = 0.2907.$$

Problem 1.3. Explain how you would sample a biased flip of a coin with probability of heads p using a uniform random variable.

Solution 1.3. If X is the outcome of a biased flip of a coin with probability of heads p, then $X \sim \mathrm{Bern}(p)$. This means that $f_X(1) = p$ and $f_X(0) = 1 - p$. The CDF and quantile function is therefore,

$$F_X(x) = \begin{cases} 0 & x < 0 \\ 1-p & 0 \leq x < 1 \\ 1 & 1 \leq x \end{cases} \qquad F_X^{-1}(x) = \begin{cases} 0 & 0 < x \leq 1-p \\ 1 & 1-p < x \leq 1 \end{cases}.$$

If $U \sim \mathrm{U}[0,1]$, we know that X has the same distribution as $F_X^{-1}(U)$. Therefore, to generate a biased coin flip, we sample $u \sim \mathrm{U}[0,1]$ and define $x(u) = 0$ if $u < 1-p$ and $x(u) = 1$ if $u > 1-p$.

1.2.2 Derivations and Proofs

Problem 1.4. Prove Theorem 1.

Solution 1.4. Let F_Y denote the CDF of the random variable $Y = F_X^{-1}(U)$. Then,

$$F_Y(x) = \mathbb{P}(F_X^{-1}(U) \le x) = \mathbb{P}(F_X(F_X^{-1}(U)) \le F_X(x)) = \mathbb{P}(U \le F(x)).$$

Furthermore, if $U \sim U(0,1)$ then

$$F_U(u) = \mathbb{P}(U \le u) = \int_0^u t\,dt = u$$

so

$$F_Y(x) = \mathbb{P}(U \le F_X(x)) = F_X(x).$$

The random variable $Y = F_X^{-1}(U)$ has the CDF F_X, as desired.

Problem 1.5. ($\star$) If F_X is a CDF, then its quantile function F_X^{-1} satisfies

$$F_X^{-1}(p) \le x \Leftrightarrow p \le F_X(x)$$

Solution 1.5. This is taken from last week's lecture summary. The proof relies on the fact that $F_X^{-1}(p)$ is the infimum of all $\{t : F_X(t) \ge p\}$, and therefore smaller than (or equal to) any $x \in \{t : F_X(t) \ge p\}$.

($\implies$) Suppose that $F_X^{-1}(p) \le x$. This implies that $x \in \{t : F_X(t) \ge p\}$ so $p \le F_X(x)$.

($\impliedby$) Suppose that $p \le F_X(x)$. This implies that $x \in \{t : F_X(t) \ge p\}$ so $F_X^{-1}(p) \le x$.

Problem 1.6. ($\star$) Prove Theorem 2.

Solution 1.6. Let F_Y denote the CDF of the random variable $Y = F_X^{-1}(U)$. We have using the properties of the quantile function (Problem 1.5) that

$$F_X^{-1}(p) \le x \Leftrightarrow p \le F_X(x).$$

So we can conclude that

$$\{F_X^{-1}(U) \le x\} = \{U \le F_X(x)\}$$

Therefore, the CDF of Y is

$$F_Y(x) = \mathbb{P}(F_X^{-1}(U) \le x) = \mathbb{P}(U \le F_X(x)) = F_X(x).$$

2 Important Continuous Random Variables Part II

2.1 Normal Distribution: $\mathrm{N}(\mu, \sigma^2)$

The normal distribution, one of the most widely and most important continuous distributions in theory and applications. A lot of data is unimodal, symmetric around the mean μ, and the majority of data is "near μ" and few are "far from μ".

Definition 1. We say that X has a *normal distribution* (or Gaussian distribution) with mean μ and variance σ^2 if X has PDF

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}, \quad x \in \mathbb{R}$$

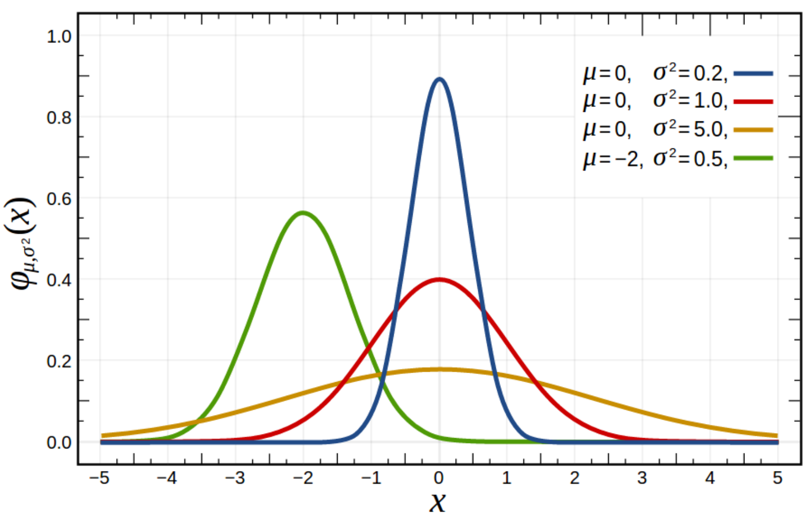
This is denoted by

$$X \sim \mathrm{N}(\mu, \sigma^2).$$

Example 3. The following experiments can be modeled by a continuous uniform distribution:

Experiment	X	Distribution
Exams with 70% average and standard dev 20%	the grade of a random student	$\mathrm{N}(0.7, 0.2^2)$
IQ test (average IQ of 100 and 15 standard dev)	IQ of a random person	$\mathrm{N}(100, 15^2)$

Example 4. Several PDFs of normally distributed random variables are below. The mean is where the peak of the PDF is, and the variance encodes how spread out the PDF is.



2.1.1 Properties

1. **Mean and Variance:** If $X \sim \mathrm{N}(\mu, \sigma^2)$ then
$$\mathbb{E}[X] = \mu \qquad \mathrm{Var}(X) = \sigma^2$$
2. **Symmetric about its mean:** If $X \sim \mathrm{N}(\mu, \sigma^2)$
$$\mathbb{P}(X \le \mu - t) = \mathbb{P}(X \ge \mu + t).$$
3. **Density is unimodal:** Peak is at μ.
4. **68-95-99.7 Rule:** If $X \sim \mathrm{N}(\mu, \sigma^2)$, then

$$\mathbb{P}(\mu - \sigma \le X \le \mu + \sigma) \approx 0.68 \qquad \text{68\% values lie within 1 standard deviation of the mean}$$
$$\mathbb{P}(\mu - 2\sigma \le X \le \mu + 2\sigma) \approx 0.95 \qquad \text{95\% values lie within 2 standard deviation of the mean}$$
$$\mathbb{P}(\mu - 3\sigma \le X \le \mu + 3\sigma) \approx 0.997 \qquad \text{99.7\% values lie within 3 standard deviation of the mean}.$$

2.2 Standard Normal Distribution

Definition 2. We say that Z follows the standard normal distribution if $Z \sim \mathrm{N}(0, 1)$. The PDF of the standard normal random variable is denoted

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}},$$

and the CDF of a standard normal random variable is denoted

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{\frac{-y^2}{2}} dy.$$

The values of $\Phi(z)$ for $z \geq 0$ and $\Phi^{-1}(p)$ for $p \geq \frac{1}{2}$ are given in $Z-$tables on the last page of the textbook. These tables are enough to compute all probabilities associated with normally distributed random variables by a technique called standardization.

Theorem 3 (*Standardising normal random variable*)

If $X \sim \mathrm{N}(\mu, \sigma^2)$, then

$$Z = \frac{X - \mu}{\sigma} \sim \mathrm{N}(0, 1).$$

Remark 2. We use the transformation $Z = \frac{X-\mu}{\sigma}$ to go from X to a standard normal. And we use the transformation $X = \mu + \sigma Z$ to go from a standard normal to X.

2.2.1 Recipe for Computing Probabilities of Normally Distributed Random Variables

CDF: To compute $\mathbb{P}(X \leq x)$ for $X \sim N(\mu, \sigma^2)$: We use the formula

$$F_X(x) = F_Z\Big(\frac{x - \mu}{\sigma}\Big).$$

To use the table (which only has half the values), we use the formula

$$F_Z(z) = \begin{cases} \Phi(z) & z \geq 0 \\ 1 - \Phi(-z) & z \leq 0 \end{cases}$$

Quantile: To compute $F_X^{-1}(p)$ for $X \sim N(\mu, \sigma^2)$: We use the formula

$$F_X^{-1}(p) = \mu + \sigma F_Z^{-1}(p).$$

To use the table (which only has half the values), we use the formula

$$F_Z^{-1}(p) = \begin{cases} \Phi^{-1}(p) & p \geq \frac{1}{2} \\ -\Phi^{-1}(1-p) & p \leq \frac{1}{2} \end{cases}$$

2.3 Example Problems

2.3.1 Applications

Problem 2.1. Suppose that $Z \sim \mathrm{N}(0, 1)$. Compute $\mathbb{P}(Z \leq -1.29)$.

Solution 2.1. We know, by symmetry of $Z \sim N(0,1)$ around $\mu = 0$, that

$$\mathbb{P}(Z \le -1.29) = \mathbb{P}(Z \ge 1.29) = 1 - \mathbb{P}(Z \le 1.29) = 1 - \Phi(1.29)$$

The table indicates $\Phi(1.29) = 0.90147$ so $\mathbb{P}(Z \le -1.29) = 1 - 0.90147 = 0.09853$

Problem 2.2. Let $X \sim N(4,2)$. Compute the following:

- $\mathbb{P}(X < 2.5)$
- $\mathbb{P}(X > 1)$
- $\mathbb{P}(-3 \le X \le 1)$
- $\mathbb{P}(X \le 4)$

Solution 2.2. Using the recipe to compute normally distributed random variables,

- $$\mathbb{P}(X < 2.5) = P(X \le 2.5) = \mathbb{P}\left(\frac{X-4}{\sqrt{2}} \le \frac{2.5-4}{\sqrt{2}}\right) = F_Z(-1.0607) = 1-\Phi(1.0607) = 1-0.85543$$
- $$\mathbb{P}(X > 1) = \mathbb{P}\left(\frac{X-4}{\sqrt{2}} > \frac{1-4}{\sqrt{2}}\right) = \mathbb{P}(Z > -2.1213) = \mathbb{P}(Z \le 2.1213) = \Phi(2.1213) = 0.9830$$
- $$\begin{aligned}\mathbb{P}(-3 \le X \le 1) &= \mathbb{P}\left(\frac{-3-4}{\sqrt{2}} \le Z \le \frac{1-4}{\sqrt{2}}\right) = \mathbb{P}(Z \le -2.1213) - \mathbb{P}(Z \le -4.95)\\ &= F_Z(-2.1213) - F_Z(-4.95) = \Phi(4.95) - \Phi(2.1213) = 1 - 0.9830 = 0.017\end{aligned}$$
- $$\mathbb{P}(X \le 4) = \mathbb{P}\left(\frac{X-4}{\sqrt{2}} \le 0\right) = \mathbb{P}(Z \le 0) = \Phi(0) = 0.5$$

Problem 2.3.

1. 75th percentile of the standard normal distribution
2. 58th percentile of the $N(5,9)$ distribution
3. Let $Z \sim N(0,1)$. Find c such that
$$\mathbb{P}(-c \le Z \le c) = 0.95$$

Solution 2.3.

1. We find
$$\Phi^{-1}(0.75) = 0.6745$$
2. We need to find the 0.58 quantile of X where $\mu = 5$ and $\sigma = \sqrt{9} = 3$,
$$F_X^{-1}(0.58) = 5 + 3F_Z^{-1}(0.58) = 5 + 3\Phi^{-1}(0.58) = 5 + 3 \cdot 0.2019 = 5.6057$$
3. We solve for c using the quantile function,
$$\begin{aligned}0.95 &= \mathbb{P}(-c \le Z \le c) = \Phi(c) - \Phi(-c)\\ &\Leftrightarrow 0.95 = \Phi(c) - (1 - \Phi(c)) = 2\Phi(c) - 1\\ &\Leftrightarrow 0.975 = \Phi(c)\\ &\Leftrightarrow c = \Phi^{-1}(0.975) = 1.96\end{aligned}$$

2.3.2 Deviations and Proofs

Problem 2.4. Prove Theorem 3

Solution 2.4. We have

$$F_Z(z) = \mathbb{P}(Z \le z) = \mathbb{P}\left(\frac{X-\mu}{\sigma} \le z\right) = \mathbb{P}(X \le \sigma z + \mu).$$

By the chain rule, for all $z \in \mathbb{R}$,

$$F_Z'(z) = \frac{d}{dz} F_X(\sigma z + \mu) = \sigma f_X(\sigma z + \mu) = \frac{\sigma}{\sqrt{2\pi\sigma^2}} e^{\frac{-(\sigma z+\mu-\mu)^2}{2\sigma^2}} = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}} = \varphi(x)$$

which we recognize as the PDF of the standard normal distribution. In particular, we have shown that $Z = \frac{X-\mu}{\sigma}$ has standard normal distribution.

Problem 2.5. If $X \sim N(\mu, \sigma^2)$ and $Z \sim N(0,1)$, show that

$$F_X(x) = F_Z\Big(\frac{x-\mu}{\sigma}\Big).$$

Solution 2.5. Using the standardization trick, we have

$$F_X(x) = \mathbb{P}(X \le x) = \mathbb{P}\left(\frac{X-\mu}{\sigma} \le \frac{x-\mu}{\sigma}\right) = \mathbb{P}\left(Z \le \frac{x-\mu}{\sigma}\right) = F_Z\Big(\frac{x-\mu}{\sigma}\Big).$$

Problem 2.6. Show that $\Phi(z) = 1 - \Phi(-z)$ for all z.

Solution 2.6. We have for all z that

$$\Phi(z) = \mathbb{P}(Z \le z) = 1 - \mathbb{P}(Z > z) = 1 - \mathbb{P}(Z \le -z) = 1 - \Phi(-z).$$

Problem 2.7. Let $Z \sim N(0,1)$ and $X \sim N(\mu, \sigma^2)$, show that $F_X^{-1}(p) = \mu + \sigma F_Z^{-1}(p)$

Solution 2.7. Let $x_p = F_X^{-1}(p)$ and $z_p = F_Z^{-1}(p)$. We have by the standardization trick,

$$p = \mathbb{P}(X \le x_p) = \mathbb{P}\left(\frac{X-\mu}{\sigma} \le \frac{x_p-\mu}{\sigma}\right) = \mathbb{P}\left(Z \le \frac{x_p-\mu}{\sigma}\right) = F_Z\left(\frac{x_p-\mu}{\sigma}\right)$$

Therefore, we can take the inverse F_Z^{-1} to see that $z_p = F_Z^{-1}(p) = \frac{x_p-\mu}{\sigma}$, so

$$x_p = \mu + \sigma z_p.$$

Problem 2.8. Show that $\Phi^{-1}(p) = -\Phi^{-1}(1-p)$ for all p.

Solution 2.8. Since the CDF is strictly increasing, we have $\Phi_Z^{-1}(p)$ is the usual inverse of a function. Let $z_p = \Phi^{-1}(p)$. Notice that by symmetry,

$$\mathbb{P}(Z \le z_p) = p \iff 1 - \mathbb{P}(Z \le z_p) = 1 - p \iff \mathbb{P}(Z > z_p) = 1 - p \iff \mathbb{P}(Z \le -z_p) = 1 - p.$$

The last equality implies that

$$\mathbb{P}(Z < -z_p) = \Phi(-z_p) = 1 - p \implies z_p = -\Phi^{-1}(1-p).$$

Problem 2.9. If $X \sim \mathrm{N}(\mu, \sigma^2)$ show that

$$\mathbb{E}[X] = \mu \qquad \mathrm{Var}(X) = \sigma^2$$

Solution 2.9. We first do the computations when $Z \sim \mathrm{N}(0,1)$. We have

$$\mathbb{E}[Z] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x e^{\frac{-x^2}{2}} = 0$$

because $xe^{\frac{-x^2}{2}}$ is an odd function. Next, to compute the second moment, we can integrate by parts

$$\mathbb{E}[Z^2] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 e^{\frac{-x^2}{2}} = -\frac{x}{\sqrt{2\pi}} e^{\frac{-x^2}{2}} \Big|_{-\infty}^{\infty} + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{\frac{-x^2}{2}} \, dx = 1$$

where the boundary term vanishes because $e^{\frac{-x^2}{2}}$ goes to zero faster than x and the second term is the integral of the PDF of a standard normal. Hence we have shown that

$$\mathbb{E}[Z] = 0 \qquad \mathrm{Var}(Z) = \mathbb{E}[Z^2] - \mathbb{E}[Z]^2 = 1.$$

To get the mean and variance of $X \sim \mathrm{N}(\mu, \sigma^2)$, we can use the standardization trick to write $X \sim \sigma Z + \mu$. Therefore,

$$\mathbb{E}[X] = \mathbb{E}[\sigma Z + \mu] = \sigma \, \mathbb{E}[Z] + \mu = \mu$$

and

$$\mathrm{Var}(X) = \mathrm{Var}(\sigma Z + \mu) = \sigma^2 \, \mathrm{Var}(Z) = \sigma^2.$$