Quiz

Let X be a random variable with p.d.f. of the type $f(x) = cx^2$, (c is a constant), which assumes values in the interval $(0, 2)$. Find the following quantities:

(a) c,

(b) $\mathsf{E}(X)$,

(c) $\mathsf{V}(X)$,

(d) median(X).

Solution:

(a) Using normalization condition:

$$\int\limits_{-\infty}^{\infty} f(x)dx = \int\limits_{0}^{2} cx^2 dx = 1, \qquad \Longrightarrow \qquad c = \frac{1}{\int\limits_0^2 x^2 dx} = \frac{1}{\left.\frac{x^3}{3}\right|_0^2} = \frac{1}{\frac{8}{3} - 0} = \boxed{\frac{3}{8}}.$$

(b) By definition of expected value:

$$\mathsf{E}(X) = \int\limits_{-\infty}^{\infty} xf(x)dx = \int\limits_0^2 \frac{3}{8}x^3 dx = \frac{3}{8} \cdot \left.\frac{x^4}{4}\right|_0^2 = \boxed{\frac{3}{2}}.$$

(c) We need $\mathsf{E}\left(X^2\right)$ to find $\mathsf{V}(X)$:

$$\mathsf{E}\left(X^2\right) = \int\limits_{-\infty}^{\infty} x^2 f(x)dx = \int\limits_0^2 \frac{3}{8}x^4 dx = \frac{3}{8} \cdot \left.\frac{x^5}{5}\right|_0^2 = \frac{12}{5}.$$

Variance then:

$$\mathsf{V}(X) = \mathsf{E}\left(X^2\right) - \mathsf{E}(X)^2 = \frac{12}{5} - \left(\frac{3}{2}\right)^2 = \boxed{\frac{3}{20}}.$$

(d) The easiest way to find median m is following:

$$\int\limits_{-\infty}^{m} f(x)dx = \int\limits_0^m \frac{3}{8}x^2 dx = \frac{1}{2}, \qquad \Longrightarrow \qquad \frac{m^3}{8} = \frac{1}{2}, \qquad \Longrightarrow \qquad m = \boxed{\sqrt[3]{4}}.$$

Problem 1

Derive a quantile function for a variable $X \sim \text{Cauchy}(0,1)$, which has the following p.d.f. on $x \in \mathbb{R}$:

$$f(x) = \frac{1}{\pi} \cdot \frac{1}{1+x^2}.$$

Solution:

Let's find c.d.f. first:

$$F(x) = \int\limits_{-\infty}^{x} f(\xi)d\xi = \int\limits_{-\infty}^{x} \frac{1}{\pi} \cdot \frac{1}{1+\xi^2} d\xi = \frac{1}{\pi} \arctan \xi \Big|_{-\infty}^{x} = \frac{1}{\pi} \arctan x - \frac{1}{\pi} \cdot \left(-\frac{\pi}{2}\right) =$$

$$= \frac{1}{\pi} \arctan x + \frac{1}{2}.$$

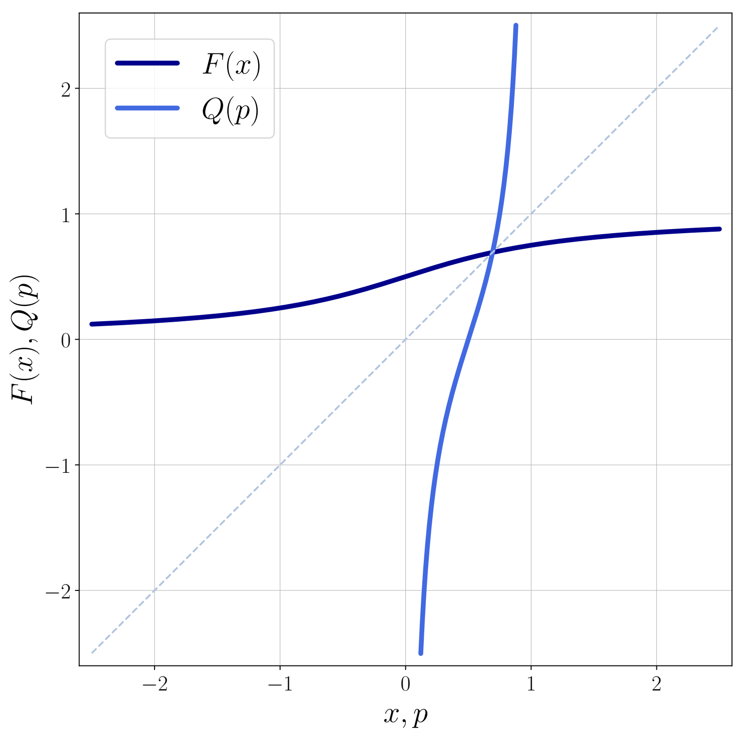

Figure 1: C.d.f. $F(x) = \dfrac{1}{\pi} \arctan x + \dfrac{1}{2}$ and quantile function $Q(p) = -\cot(\pi p)$.

$F(x)$ is a strictly monotonic function, which has inverse $Q(p)$.

$$\begin{aligned}
\frac{1}{\pi}\arctan Q(p)+\frac{1}{2}&=p,\\
\arctan Q(p)&=\pi p-\frac{\pi}{2},\\
Q(p)&=\tan\left(\pi p-\frac{\pi}{2}\right)=\boxed{-\cot(\pi p)},
\end{aligned}$$

where p takes values from interval $[0,1]$.

The symmetry over the 1$^{\text{st}}$ quadrant bisector between functions $F(x)$ and $Q(p)$ is illustrated in the fig. 1.

Problem 2

Derive and sketch a quantile function for $X \sim \text{Bernoulli}\left(\frac{1}{2}\right)$.

Solution:

X has following p.m.f.:

X	0	1
P_X	$\frac{1}{2}$	$\frac{1}{2}$

Which means that a c.d.f. $F(x)$ of X has the view:

$$F(x) = \begin{cases} 1, & x \geq 1, \\ \frac{1}{2}, & 0 \leq x < 1, \\ 0, & x < 0, \end{cases}$$

and illustrated in the fig. 2.

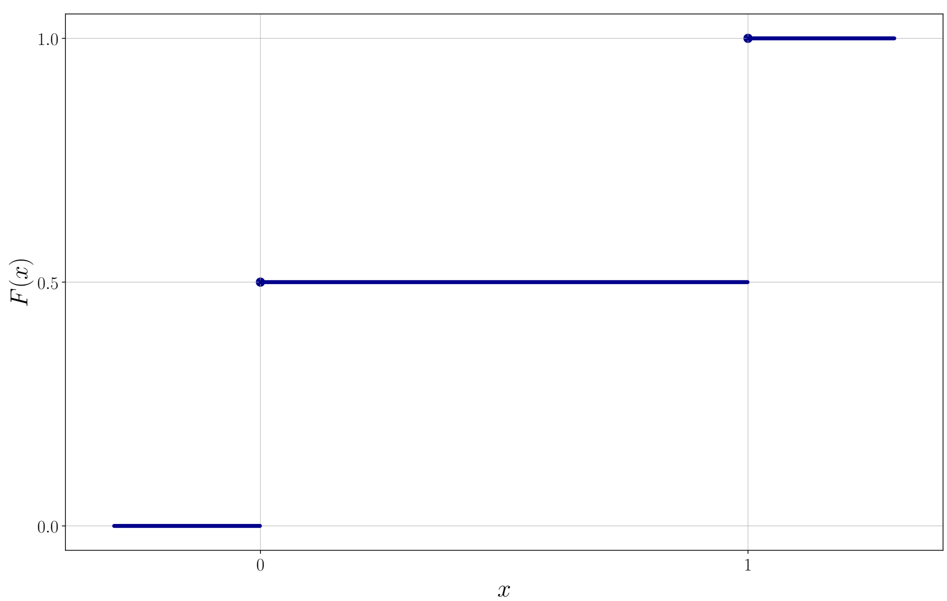

Figure 2: C.d.f. $F(x)$ of $X \sim \text{Bernoulli}\left(\frac{1}{2}\right)$.

According to the definition of $Q(p)$, which takes into account right-continuity of $F(x)$

$$Q(p) = \inf\{x \in \mathbb{R} : p \leq F_X(x)\},$$

and attaching the value from $Q(0) = -\infty$ to $Q(0) = 0$ (for the convenience of usage):

$$Q(p) = \begin{cases} 1, & \dfrac{1}{2} < x \leq 1, \\ 0, & 0 \leq x \leq \dfrac{1}{2}. \end{cases}$$

$Q(p)$ is illustrated in the fig. 3.

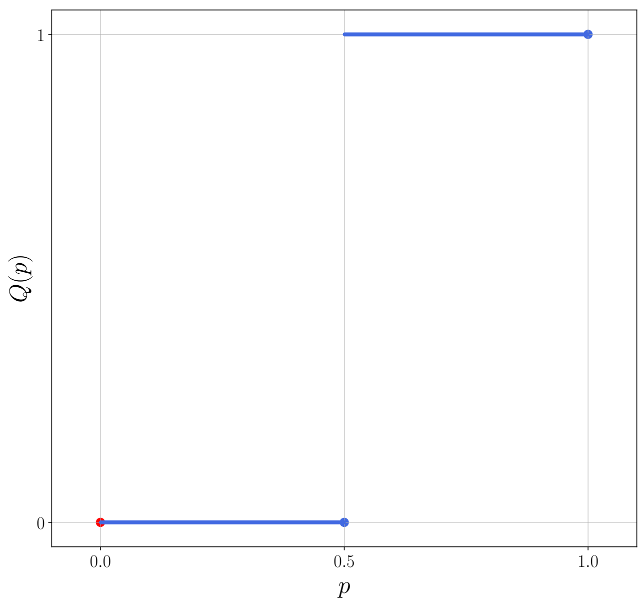

Figure 3: Quantile function $Q(p)$ of $X \sim \text{Bernoulli}\left(\dfrac{1}{2}\right)$.

Problem 3

Consider normally distributed random variable $Z \sim \mathcal{N}(0, 1^2)$.

(a) What is $\mathsf{P}(Z > 1.2)$?

(b) What is $\mathsf{P}(-1.24 \le Z \le 1.86)$?

Solution:

(a) Using total probability:

$$\mathsf{P}(Z > 1.2) = 1 - \mathsf{P}(Z \le 1.2),$$

which by definition of c.d.f. is:

$$\mathsf{P}(Z > 1.2) = 1 - \Phi(1.2),$$

where $\Phi(z)$ is a c.d.f. of standard normal distribution $\mathcal{N}(0, 1)$. Values of $\Phi(z)$ are found via standard normal distribution table:

$$\mathsf{P}(Z > 1.2) = 1 - \Phi(1.2) \approx 1 - 0.885 = \boxed{0.115}.$$

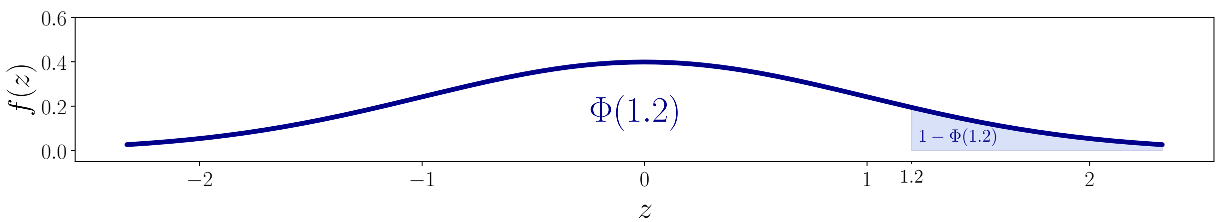

Figure 4: P.d.f. $f(z)$ with highlighted $\mathsf{P}(Z > 1.2)$.

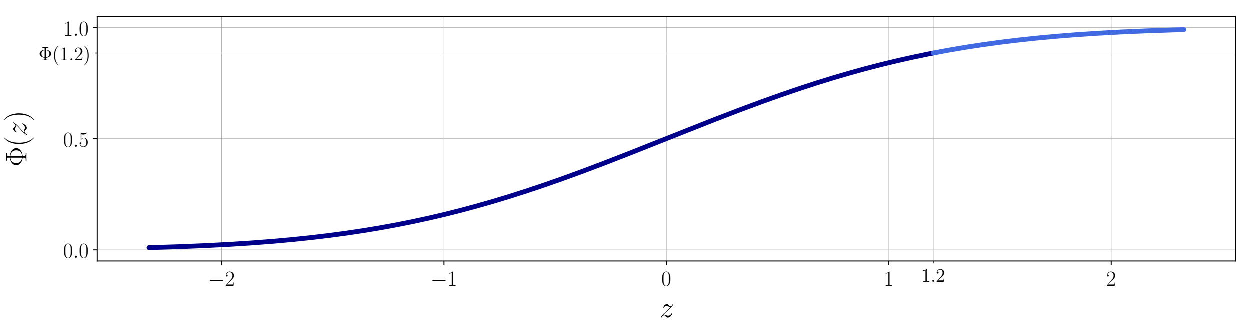

Figure 5: C.d.f. $\Phi(z)$ with highlighted $\mathsf{P}(Z > 1.2)$.

Geometric interpretation of $\mathsf{P}(Z > 1.2)$ is shown in fig. 4 and 5.

(b) Probability in the interval is found via difference of respective c.d.f.s:

$$\mathsf{P}(-1.24 \leq Z \leq 1.86) = \Phi(1.86) - \Phi(-1.24),$$

which has clear geometric interpretation (see fig. 6 and 7).

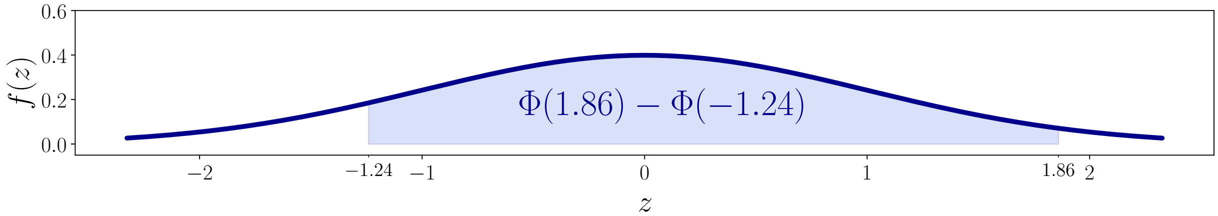

Figure 6: P.d.f. $f(z)$ with highlighted $\mathsf{P}(-1.24 \leq Z \leq 1.86)$.

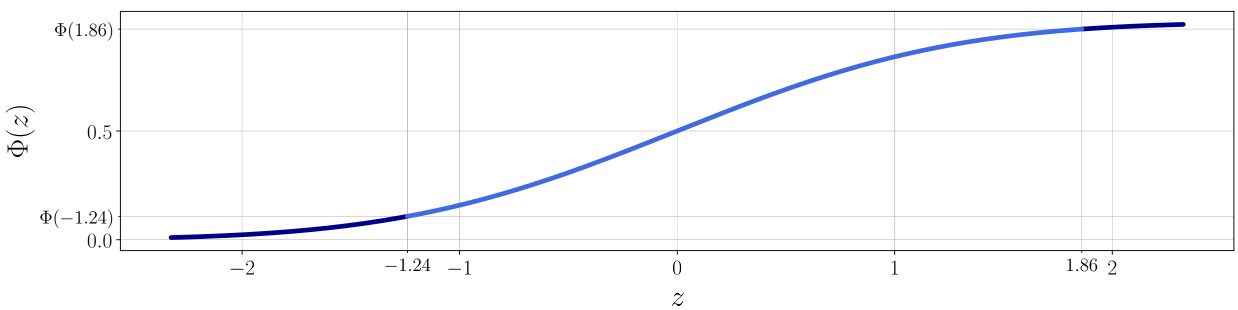

Figure 7: C.d.f. $\Phi(z)$ with highlighted $\mathsf{P}(-1.24 \leq Z \leq 1.86)$.

There could be a problem with $\Phi(-1.24)$, since standard normal distribution tables usually do not have negative z-values. We should use the symmetry of Z, which results in following property (equal areas of equidistant tails of p.d.f. $f(z)$):

$$\Phi(-z) = 1 - \Phi(z).$$

Using the symmetry:

$$\begin{aligned}\mathsf{P}(-1.24 \leq Z \leq 1.86) &= \Phi(1.86) - \Phi(-1.24) = \\ &= \Phi(1.86) - 1 + \Phi(1.24) \approx 0.969 - 1 + 0.892 = \boxed{0.861}.\end{aligned}$$

Problem 4

Consider normally distributed random variable $X \sim \mathcal{N}(5, 4)$.

(a) What is $\mathsf{P}(X > 6.4)$?

(b) What is $\mathsf{P}(5.8 \leq X < 7.0)$?

Solution:

(a) Since we do not have table of values for c.d.f. of $X \sim \mathcal{N}(5, 4)$, we have to standardize X to standard normal variable $Z \sim \mathcal{N}(0, 1)$. Standardizing normal variable X with mean μ and variance σ^2 is done as follows:

$$Z = \frac{X - \mu}{\sigma}.$$

It means that required probability can be rewritten via Z. Applying mentioned standardization to all pars of inequality:

$$\mathsf{P}(X > 6.4) = \mathsf{P}\left(\frac{X-5}{2} > \frac{6.4-5}{2}\right) = \mathsf{P}(Z > 0.7).$$

Using total probability, definition of c.d.f. and standard normal distribution table:

$$\mathsf{P}(X > 6.4) = \mathsf{P}(Z > 0.7) = 1 - \mathsf{P}(Z \leq 0.7) = 1 - \Phi(0.7) \approx 1 - 0.758 = \boxed{0.242}.$$

Standardization of $\mathsf{P}(X > 6.4)$ is shown in fig. 8.

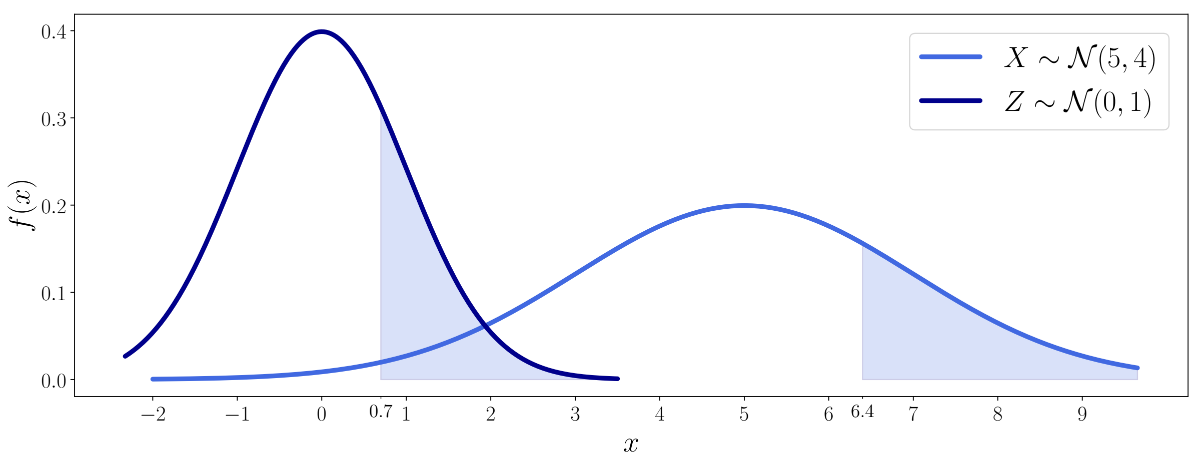

Figure 8: Equivalence of $\mathsf{P}(X > 6.4)$ and $\mathsf{P}(Z > 0.7)$.

(b) Applying standardization to all 3 parts of inequality gives:

$$\mathsf{P}(5.8 \leq X < 7.0) = \mathsf{P}\left(\frac{5.8-5}{2} \leq \frac{X-5}{2} < \frac{7.0-5}{2}\right) = \mathsf{P}(0.4 \leq Z < 1.0).$$

Using difference of c.d.f.s and standard normal distribution table:

$$\mathsf{P}(5.8 \leq X < 7.0) = \mathsf{P}(0.4 \leq Z < 1.0) = \Phi(1.0) - \Phi(0.4) \approx 0.841 - 0.655 = \boxed{0.186}.$$

Standardization of $\mathsf{P}(5.8 \leq X < 7.0)$ is shown in fig. 9.

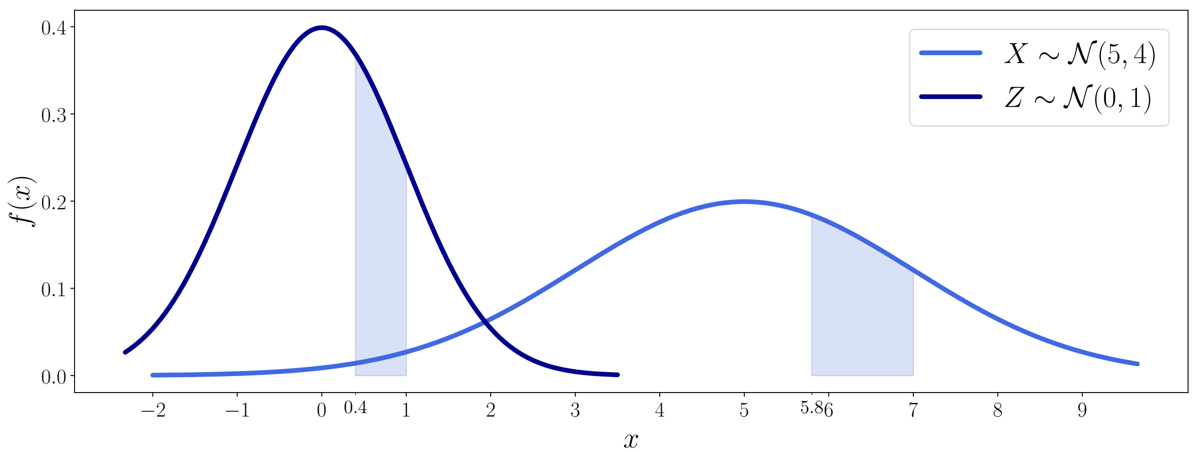

Figure 9: Equivalence of $\mathsf{P}(5.8 \leq X < 7.0)$ and $\mathsf{P}(0.4 \leq Z < 1.0)$.

Problem 5

Consider normally distributed random variable $X \sim \mathcal{N}(6, 25)$.

(a) Find $\mathsf{P}(6 < X < 12)$.

(b) Find $\mathsf{P}(0 \leq X < 8)$.

(c) Find $\mathsf{P}(-2 < X \leq 0)$.

Solution:

(a) Applying standardization:

$$\mathsf{P}(6 < X < 12) = \mathsf{P}\left(\frac{6-6}{5} < \frac{X-6}{5} < \frac{12-6}{5}\right) = \mathsf{P}(0 < Z < 1.2).$$

Using difference of c.d.f.s and standard normal distribution table:

$$\mathsf{P}(6 < X < 12) = \mathsf{P}(0 < Z < 1.2) = \Phi(1.2) - \Phi(0) \approx 0.885 - 0.5 = \boxed{0.385}.$$

Standardizations are shown in fig. 10, 11 and 12.

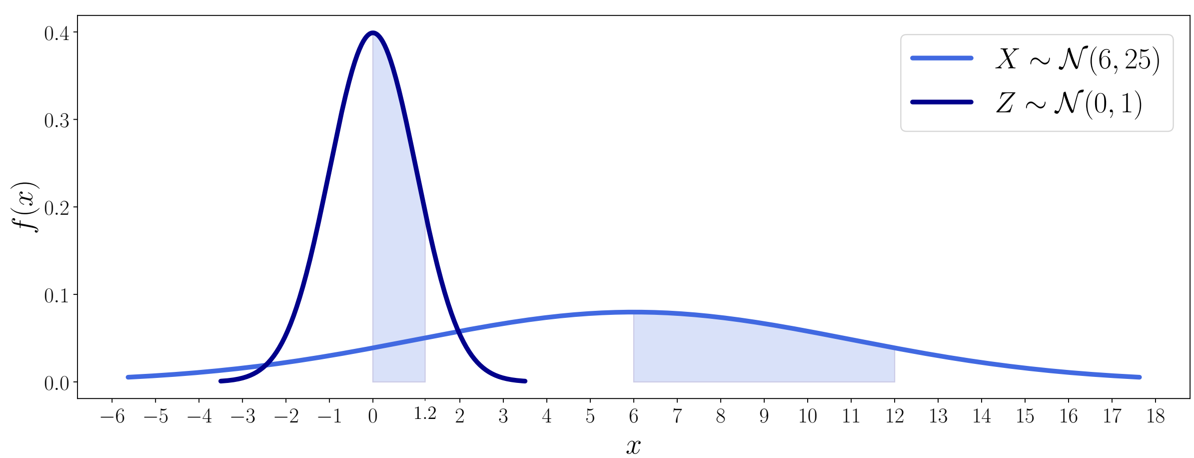

Figure 10: Equivalence of $\mathsf{P}(6 < X < 12)$ and $\mathsf{P}(0 < Z < 1.2)$.

(b) Applying standardization:

$$\mathsf{P}(0 \leq X < 8) = \mathsf{P}\left(\frac{0-6}{5} \leq \frac{X-6}{5} < \frac{8-6}{5}\right) = \mathsf{P}(-1.2 \leq Z < 0.4).$$

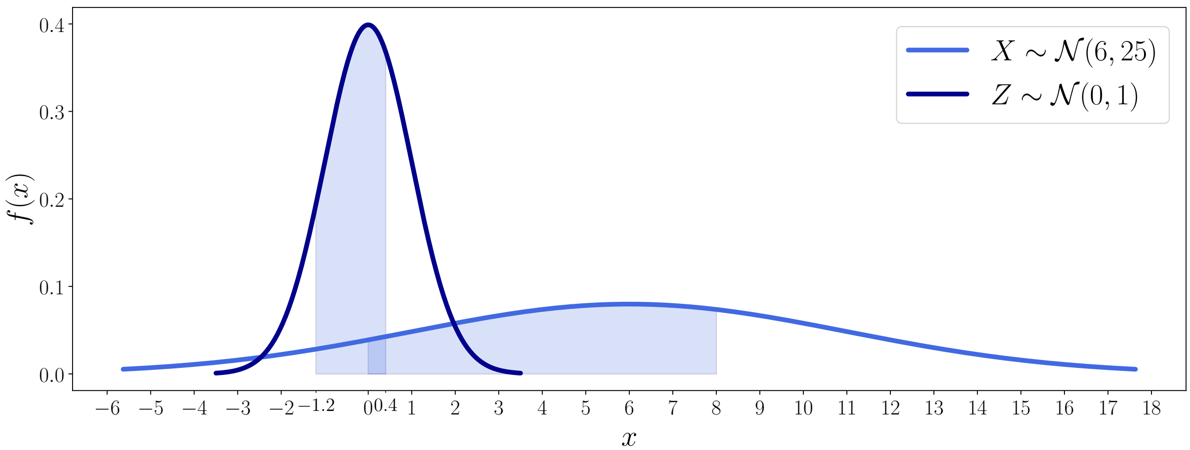

Figure 11: Equivalence of $\mathsf{P}(0 \le X < 8)$ and $\mathsf{P}(-1.2 \le Z < 0.4)$.

Using difference of c.d.f.s, symmetry of Z and standard normal distribution table:

$$\mathsf{P}(0 \le X < 8) = \mathsf{P}(-1.2 \le Z < 0.4) = \Phi(0.4) - \Phi(-1.2) =$$
$$= \Phi(0.4) - 1 + \Phi(1.2) \approx 0.655 - 1 + 0.885 = \boxed{0.540}.$$

(c) Repeating previous item:

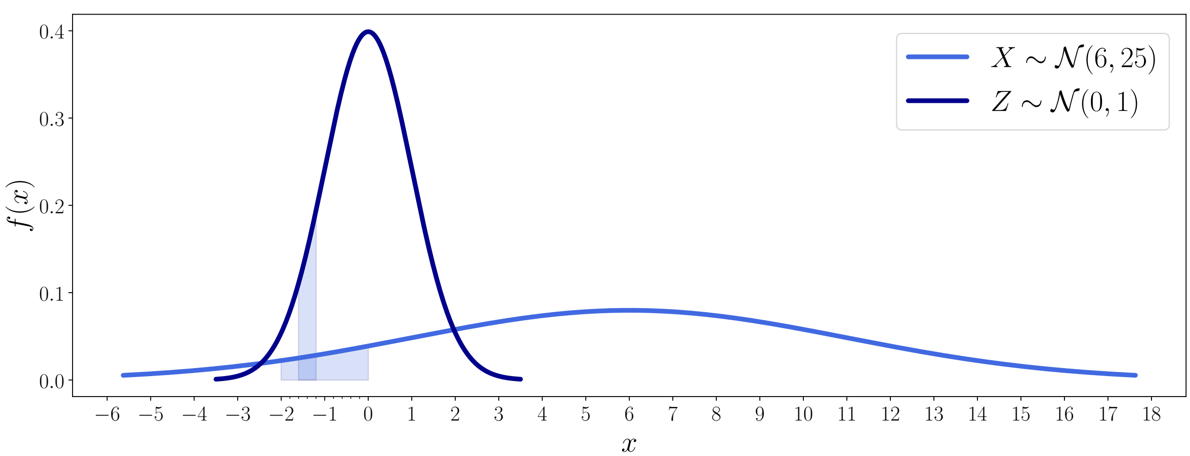

Figure 12: Equivalence of $\mathsf{P}(-2 < X \le 0)$ and $\mathsf{P}(-1.6 < Z \le -1.2)$.

$$\mathsf{P}(-2 < X \le 0) = \mathsf{P}\left(\frac{-2-6}{5} < \frac{X-6}{5} \le \frac{0-6}{5}\right) = \mathsf{P}(-1.6 < Z \le -1.2).$$

$$\mathsf{P}(-2 < X \le 0) = \mathsf{P}(-1.6 < Z \le -1.2) = \Phi(-1.2) - \Phi(-1.6) =$$
$$= \Phi(1.6) - \Phi(1.2) \approx 0.945 - 0.885 = \boxed{0.060}.$$

Problem 6

The manufacturer of a brand new lithium battery claims that the mean life of a battery is 3800 hours with a standard deviation of 250 hours.

(a) What percentage of batteries will last for more than 3500 hours?

(b) What percentage of batteries will last for more than 4000 hours?

(c) Batteries, which will last for more than c hours, constitute more than $1/5$ of the population. Find the maximum possible c.

Solution:

Let's assume that mean lifetime of batteries is distributed normally with parameters: $X \sim \mathcal{N}(3800, 250^2)$ in hours.

(a) Using standardization, symmetry of Z and standard normal distribution tables:

$$\mathsf{P}(X > 3500) = \mathsf{P}\left(\frac{X - 3800}{250} > \frac{3500 - 3800}{250}\right) = \mathsf{P}(Z > -1.2) =$$
$$= 1 - \Phi(-1.2) = \Phi(1.2) \approx \boxed{0.885}.$$

(b) Using standardization, total probability and standard normal distribution tables:

$$\mathsf{P}(X > 4000) = \mathsf{P}\left(\frac{X - 3800}{250} > \frac{4000 - 3800}{250}\right) = \mathsf{P}(Z > 0.8) =$$
$$= 1 - \Phi(0.2) \approx 1 - 0.788 = \boxed{0.212}.$$

(c) We want to find maximal possible c such that the following inequality holds:

$$\mathsf{P}(X > c) \geq \frac{1}{5}.$$

Using standardization, total probability and standard normal distribution tables:

$$\mathsf{P}(X > c) = \mathsf{P}\left(Z > \frac{c - 3800}{250}\right) = 1 - \Phi\left(\frac{c - 3800}{250}\right) \geq \frac{1}{5},$$

$$\Phi\left(\frac{c - 3800}{250}\right) \leq \frac{4}{5}.$$

The closest value from table, which satisfies previous inequality is 0.84:

$$\frac{c - 3800}{250} \approx 0.84 \qquad \Longrightarrow \qquad c \approx 250 \cdot 0.84 + 3800 = \boxed{4010}.$$

Direct calculation in Python gives: $c = 4010.405$ hours.

Problem 7

Two manufacturers of lithium batteries produce them with different specifications: the mean life of a battery from the first manufacturer is 3800 hours with a standard deviation of 250 hours, the second one's mean life is 3600 hours with standard deviation of 280 hours.

Suppose we take one battery of each kind at random and independently from each other. What is the probability that the lifetime of the first battery will be at most 50 hours greater than the second's one?

Solution:

Let's assume that lifetimes of both batteries types are distributed normally with respective parameters: $X_1 \sim \mathcal{N}\left(3800, 250^2\right)$ and $X_2 \sim \mathcal{N}\left(3600, 280^2\right)$ in hours. We are asked to calculate the probability $\mathsf{P}(X_1 - X_2 < 50)$.

The linear combination of n normal random variables $X = \sum\limits_{i=1}^{n} a_i X_i + b$, where a_i and b are constants and $X_i \sim \mathcal{N}\left(\mu_i, \sigma_i^2\right)$:

- is also normal $X \sim \mathcal{N}\left(\mu, \sigma^2\right)$,
- with mean $\mu = \sum\limits_{i=1}^{n} a_i \mu_i + b$,
- and variance $\sigma^2 = \sum\limits_{i=1}^{n} a_i^2 \sigma_i^2 + 2 \sum\limits_{i<j}\sum a_i a_j \mathrm{Cov}(X_i, X_j)$. If all X_i are independent, then $\mathrm{Cov}(X_i, X_j) = 0$ and the variance of their combination is simply $\sigma^2 = \sum\limits_{i=1}^{n} a_i^2 \sigma_i^2$.

Let random variable D be the difference between X_1 and X_2: $D = X_1 - X_2$. Since X_1 and X_2 are independent:

$$D \sim \mathcal{N}\left(3800 - 3600, 250^2 + 280^2\right) \approx \mathcal{N}\left(200, 375.37^2\right).$$

Now we need to find the probability $\mathsf{P}(D < 50)$. Standardizing the random variable D and limit to use standard normal distribution table of c.d.f. $\Phi(z)$:

$$\mathsf{P}(D < 50) = \mathsf{P}\left(\frac{D - 200}{375.37} < \frac{50 - 200}{375.37}\right) = \mathsf{P}(Z < -0.3996) = \Phi(-0.3996).$$

Using the symmetry of standard normal distribution $\Phi(-z) = 1 - \Phi(z)$:

$$\mathsf{P}(X_1 - X_2 < 50) = \Phi(-0.3996) \approx 1 - \Phi(0.3996) = 1 - 0.655 = \boxed{0.345}.$$