Practice 5

Carlo Cavicchia

ccavicchia@luiss.it

Exercise 1

The length of Black Mirror episodes (in minutes), can be assumed to be Normally distributed with mean $\mu = 50$ minutes and standard deviation $\sigma = 5$ minutes. A new season has just been released.

- What is the probability that the length of an episode of the new season is exactly 50 minutes?

Exercise 1

The length of Black Mirror episodes (in minutes), can be assumed to be Normally distributed with mean $\mu = 50$ minutes and standard deviation $\sigma = 5$ minutes. A new season has just been released.

- What is the probability that the length of an episode of the new season is exactly 50 minutes?

Let T denote the episode's length and assume $T \sim N(\mu = 50, \sigma^2 = 5^2)$.

Exercise 1

The length of Black Mirror episodes (in minutes), can be assumed to be Normally distributed with mean $\mu = 50$ minutes and standard deviation $\sigma = 5$ minutes. A new season has just been released.

- What is the probability that the length of an episode of the new season is exactly 50 minutes?

Let T denote the episode's length and assume $T \sim N(\mu = 50, \sigma^2 = 5^2)$.

Any interval with length 0 has null probability, hence $P(T = 50) = 0$.

Exercise 1

The length of Black Mirror episodes (in minutes), can be assumed to be Normally distributed with mean $\mu = 50$ minutes and standard deviation $\sigma = 5$ minutes. A new season has just been released.

- What is the probability that the length of an episode of the new season is between 48 and 51 minutes?

Exercise 1

The length of Black Mirror episodes (in minutes), can be assumed to be Normally distributed with mean $\mu = 50$ minutes and standard deviation $\sigma = 5$ minutes. A new season has just been released.

- What is the probability that the length of an episode of the new season is between 48 and 51 minutes?

Let us start with the standardization of T:

$$Z = \frac{T - \mu}{\sigma}$$

So,

$$P(48 < T \leq 51) = P(\frac{48 - \mu}{\sigma} < \frac{T - \mu}{\sigma} \leq \frac{51 - \mu}{\sigma}) = P(-0.4 < Z \leq 0.2)$$

Exercise 1

The length of Black Mirror episodes (in minutes), can be assumed to be Normally distributed with mean $\mu = 50$ minutes and standard deviation $\sigma = 5$ minutes. A new season has just been released.

- What is the probability that the length of an episode of the new season is between 48 and 51 minutes?

$$P(48 < T \leq 51) = P(-0.4 < Z \leq 0.2) = \phi(0.2) - \phi(-0.4) =$$

$$\phi(0.2) - [1 - \phi(0.4)] = 0.5793 - 0.3446 = 0.2347$$

Exercise 1

The length of Black Mirror episodes (in minutes), can be assumed to be Normally distributed with mean $\mu = 50$ minutes and standard deviation $\sigma = 5$ minutes. A new season has just been released.

The new season, composed of 9 episodes, is scheduled to be released next fall. We are interested in the total length of the season (length of all 9 episodes played in sequence).

- What is the probability distribution of the total length?

Exercise 1

The length of Black Mirror episodes (in minutes), can be assumed to be Normally distributed with mean $\mu = 50$ minutes and standard deviation $\sigma = 5$ minutes. A new season has just been released.

The new season, composed of 9 episodes, is scheduled to be released next fall. We are interested in the total length of the season (length of all 9 episodes played in sequence).

- What is the probability distribution of the total length?

The total length $L = \sum_{i=1}^{9} T_i$ is sum of independent normal random variables, hence L is normal with expectation the sum of the single expectations and variance the sum of the single variances.

In conclusion, $L \sim N(\mu = 450, \sigma^2 = 225)$

Exercise 1

The length of Black Mirror episodes (in minutes), can be assumed to be Normally distributed with mean $\mu = 50$ minutes and standard deviation $\sigma = 5$ minutes. A new season has just been released.

The new season, composed of 9 episodes, is scheduled to be released next fall. We are interested in the total length of the season (length of all 9 episodes played in sequence).

- Now consider the total length of the series in hours (for instance, if the total length is 405 minutes, in hours it will be $405/60 = 6.75$). Determine expected value and variance of the total length in hours.

Exercise 1

The length of Black Mirror episodes (in minutes), can be assumed to be Normally distributed with mean $\mu = 50$ minutes and standard deviation $\sigma = 5$ minutes. A new season has just been released.

The new season, composed of 9 episodes, is scheduled to be released next fall. We are interested in the total length of the season (length of all 9 episodes played in sequence).

- Now consider the total length of the series in hours (for instance, if the total length is 405 minutes, in hours it will be $405/60 = 6.75$). Determine expected value and variance of the total length in hours.

The total length in hours is $H = L/60$. Hence its expectation and variance are: $\mathrm{E}(H) = \mathrm{E}(L)/60 = 7.5$ and $\mathrm{V}(H) = \mathrm{V}(L)/60^2 = 0.0625$

Exercise 1

The length of Black Mirror episodes (in minutes), can be assumed to be Normally distributed with mean $\mu = 50$ minutes and standard deviation $\sigma = 5$ minutes. A new season has just been released.

The new season, composed of 9 episodes, is scheduled to be released next fall. We are interested in the total length of the season (length of all 9 episodes played in sequence).

- Which hypotheses have been implicitely assumed in answering the last two points?

Exercise 1

The length of Black Mirror episodes (in minutes), can be assumed to be Normally distributed with mean $\mu = 50$ minutes and standard deviation $\sigma = 5$ minutes. A new season has just been released.

The new season, composed of 9 episodes, is scheduled to be released next fall. We are interested in the total length of the season (length of all 9 episodes played in sequence).

- Which hypotheses have been implicitely assumed in answering the last two points?

The episodes' lengths $T_1, \dots, T_9$ are all assumed to be normally distributed with parameters $(\mu = 50, \sigma^2 = 25)$. We also suppose that the lengths are all independent.

Exercise 2

The time to run 10km spent by an athlete is distributed as a normal $(45, 25)$, whereas the same time spent by a non-athlete is distributed as a normal $(50, 100)$. 20 athletes and 80 non-athletes take part of a 10km race.

Exercise 2

The time to run 10km spent by an athlete is distributed as a normal $(45, 25)$, whereas the same time spent by a non-athlete is distributed as a normal $(50, 100)$. 20 athletes and 80 non-athletes take part of a 10km race.

- What is the probability that a runner finishes the race with a time shorter than 50 minutes?

Exercise 2

The time to run 10km spent by an athlete is distributed as a normal $(45, 25)$, whereas the same time spent by a non-athlete is distributed as a normal $(50, 100)$. 20 athletes and 80 non-athletes take part of a 10km race.

- What is the probability that a runner finishes the race with a time shorter than 50 minutes?

First of all, let us define the RVs we are considering.

$$T|A \sim N(45, 25)$$

$$T|\bar{A} \sim N(50, 100)$$

Exercise 2

The time to run 10km spent by an athlete is distributed as a normal $(45, 25)$, whereas the same time spent by a non-athlete is distributed as a normal $(50, 100)$. 20 athletes and 80 non-athletes take part of a 10km race.

- What is the probability that a runner finishes the race with a time shorter than 50 minutes?

So, for a generic runner we have

$$P(T = t) = P(T = t|A)P(A) + P(T = t|\bar{A})P(\bar{A})$$

Exercise 2

The time to run 10km spent by an athlete is distributed as a normal $(45, 25)$, whereas the same time spent by a non-athlete is distributed as a normal $(50, 100)$. 20 athletes and 80 non-athletes take part of a 10km race.

- What is the probability that a runner finishes the race with a time shorter than 50 minutes?

$$P(T < 50) = P(T < 50|A)P(A) + P(T < 50|\bar{A})P(\bar{A})$$

Exercise 2

The time to run 10km spent by an athlete is distributed as a normal $(45, 25)$, whereas the same time spent by a non-athlete is distributed as a normal $(50, 100)$. 20 athletes and 80 non-athletes take part of a 10km race.

- What is the probability that a runner finish the race with a time shorter than 50 minutes?

Let us standardize our RV for the value 50:

$$Z_A = \frac{50 - 45}{5} = 1$$

$$Z_{\bar{A}} = \frac{50 - 50}{10} = 0$$

Exercise 2

The time to run 10km spent by an athlete is distributed as a normal $(45, 25)$, whereas the same time spent by a non-athlete is distributed as a normal $(50, 100)$. 20 athletes and 80 non-athletes take part of a 10km race.

- What is the probability that a runner finish the race with a time shorter than 50 minutes?

Thus,

$$P(T < 50) = P(T < 50|A)P(A) + P(T < 50|\bar{A})P(\bar{A})$$

Exercise 2

The time to run 10km spent by an athlete is distributed as a normal $(45, 25)$, whereas the same time spent by a non-athlete is distributed as a normal $(50, 100)$. 20 athletes and 80 non-athletes take part of a 10km race.

- What is the probability that a runner finish the race with a time shorter than 50 minutes?

Thus,

$$P(T < 50) = P(Z < Z_A)P(A) + P(Z < Z_{\bar{A}})P(\bar{A}) = P(Z < 1)\frac{20}{100} + P(Z < 0)\frac{80}{100}$$

$$\frac{1}{5}\phi(1) + \frac{4}{5}\frac{1}{2} = 0.2 \times 0.8413 + 0.8 \times 0.5 \approx 0.57$$

Exercise 2

The time to run 10km spent by an athlete is distributed as a normal $(45, 25)$, whereas the same time spent by a non-athlete is distributed as a normal $(50, 100)$. 20 athletes and 80 non-athletes take part of a 10km race.

Exercise 2

The time to run 10km spent by an athlete is distributed as a normal $(45, 25)$, whereas the same time spent by a non-athlete is distributed as a normal $(50, 100)$. 20 athletes and 80 non-athletes take part of a 10km race.

- A runner arrives after 50 minutes. What is the probability that he/she is an athlete?

Exercise 2

The time to run 10km spent by an athlete is distributed as a normal $(45, 25)$, whereas the same time spent by a non-athlete is distributed as a normal $(50, 100)$. 20 athletes and 80 non-athletes take part of a 10km race.

- A runner arrives after 50 minutes. What is the probability that he/she is an athlete?

$$P(A|T > 50) = \frac{P(T > 50|A)P(A)}{P(T > 50)} = \frac{P(Z > Z_A)\frac{1}{5}}{1 - P(T < 50)}$$

$$\frac{[1 - P(Z < 1)]\frac{1}{5}}{0.43} = \frac{0.1587 \times \frac{1}{5}}{0.43} \approx 0.07$$

Exercise 3

Theoretical exercise

If X is normally distributed with mean μ and variance $\sigma^2 > 0$, then:

$$V = \left(\frac{X - \mu}{\sigma}\right)^2 = Z^2$$

is distributed as a chi-square random variable with 1 degree of freedom.

Exercise 3

Theoretical exercise

To prove this theorem, we need to show that the p.d.f. of the random variable V is the same as the p.d.f. of a chi-square random variable with 1 degree of freedom. That is, we need to show that:

$$f_V(v) = \frac{1}{\Gamma\left(\frac{1}{2}\right) 2^{\frac{1}{2}}} v^{\frac{1}{2}-1} e^{-\frac{v}{2}}$$

Exercise 3

Theoretical exercise

The strategy we'll take is to find $F_V(v)$, the cumulative distribution function of V, and then differentiate it to get $f_V(v)$, the probability density function of V. That said, we start with the definition of the cumulative distribution function of V:

Exercise 3

Theoretical exercise

The strategy we'll take is to find $F_V(v)$, the cumulative distribution function of V, and then differentiate it to get $f_V(v)$, the probability density function of V. That said, we start with the definition of the cumulative distribution function of V:

$$F_V(v) = P(V \leq v) = P(Z^2 \leq v) = P(-\sqrt{v} \leq Z \leq \sqrt{v})$$

Exercise 3

Theoretical exercise

The strategy we'll take is to find $F_V(v)$, the cumulative distribution function of V, and then differentiate it to get $f_V(v)$, the probability density function of V. That said, we start with the definition of the cumulative distribution function of V:

$$F_V(v) = P(V \leq v) = P(Z^2 \leq v) = P(-\sqrt{v} \leq Z \leq \sqrt{v})$$

So,

$$F_V(v) = \int_{-\sqrt{v}}^{\sqrt{v}} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz$$

Exercise 3

Theoretical exercise

The strategy we'll take is to find $F_V(v)$, the cumulative distribution function of V, and then differentiate it to get $f_V(v)$, the probability density function of V. That said, we start with the definition of the cumulative distribution function of V:

By the symmetry of the normal distribution, we can integrate over just the positive portion of the integral, and then multiply by two:

$$F_V(v) = 2\int_0^{\sqrt{v}} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}\, dz$$

Exercise 3

Theoretical exercise

The strategy we'll take is to find $F_V(v)$, the cumulative distribution function of V, and then differentiate it to get $f_V(v)$, the probability density function of V. That said, we start with the definition of the cumulative distribution function of V:

Let us consider the following change of variables:

$$z = y^{\frac{1}{2}}$$

So:

$$dz = \frac{1}{2}y^{-\frac{1}{2}}dy = \frac{dy}{2\sqrt{y}}$$

Exercise 3

Theoretical exercise

The strategy we'll take is to find $F_V(v)$, the cumulative distribution function of V, and then differentiate it to get $f_V(v)$, the probability density function of V. That said, we start with the definition of the cumulative distribution function of V:

$$F_V(v) = 2\int_0^{\sqrt{v}} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}\,dz = 2\int_0^{v} \frac{1}{\sqrt{2\pi}} e^{-\frac{y}{2}}\,\frac{dy}{2\sqrt{y}}$$

$$F_V(v) = \int_0^{v} \frac{1}{\sqrt{\pi}\sqrt{2}} y^{\frac{1}{2}-1} e^{-\frac{y}{2}}\,dy$$

Exercise 3

Theoretical exercise

The strategy we'll take is to find $F_V(v)$, the cumulative distribution function of V, and then differentiate it to get $f_V(v)$, the probability density function of V. That said, we start with the definition of the cumulative distribution function of V:

$$f_V(v) = F'_V(v) = \frac{1}{\sqrt{\pi}\sqrt{2}} v^{\frac{1}{2}-1} e^{-\frac{v}{2}}$$

for $v > 0$.

Exercise 3

Theoretical exercise

The strategy we'll take is to find $F_V(v)$, the cumulative distribution function of V, and then differentiate it to get $f_V(v)$, the probability density function of V. That said, we start with the definition of the cumulative distribution function of V:

If you compare this $g_V(v)$ to the first $g_V(v)$ that we said we needed to find way back at the beginning of this proof, you should see that we are done if the following is true:

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$$

Exercise 3

Theoretical exercise

The strategy we'll take is to find $F_V(v)$, the cumulative distribution function of V, and then differentiate it to get $f_V(v)$, the probability density function of V. That said, we start with the definition of the cumulative distribution function of V:

It is indeed true, as the following argument illustrates. Because $g_V(v)$ is a p.d.f., the integral of the p.d.f. over the support must equal 1:

$$\int_0^\infty \frac{1}{\sqrt{\pi}\sqrt{2}} v^{\frac{1}{2}-1} e^{-\frac{v}{2}}\, dv = 1$$

Exercise 3

Theoretical exercise

The strategy we'll take is to find $F_V(v)$, the cumulative distribution function of V, and then differentiate it to get $f_V(v)$, the probability density function of V. That said, we start with the definition of the cumulative distribution function of V:

Now, change the variables by letting $v = 2x$, so that $dv = 2dx$. Making the change, we get:

$$\frac{1}{\sqrt{\pi}}\int_0^\infty \frac{1}{\sqrt{2}}(2x)^{\frac{1}{2}-1}e^{-x}2dx = \frac{1}{\sqrt{\pi}}\int_0^\infty x^{\frac{1}{2}-1}e^{-x}dx = \frac{1}{\sqrt{\pi}}\Gamma\left(\frac{1}{2}\right) = 1$$

Exercise 4

Find the probability that the standard normal random variable Z falls between -1.96 and 1.96.

- using the standard normal distribution

Exercise 4

Find the probability that the standard normal random variable Z falls between -1.96 and 1.96.

- using the standard normal distribution

$$P(-1.96 \leq Z \leq 1.96) = P(Z \leq 1.96)) - P(Z \leq -1.96)) =$$

$$P(Z \leq 1.96)) - P(Z \geq 1.96)) = 0.975 - 0.025 = 0.95$$

Exercise 4

Find the probability that the standard normal random variable Z falls between -1.96 and 1.96.

- using the chi-square distribution

Exercise 4

Find the probability that the standard normal random variable Z falls between -1.96 and 1.96.

› using the chi-square distribution

$$P(-1.96 < Z < 1.96) = P(|Z| < 1.96) = P(Z^2 < 1.96^2) = P(\chi^2_{(1)} < 3.8416) = 0.95$$

Exercise 5

First and third quartile of a Normal RV are 1 and 3, respectively.

Exercise 5

First and third quartile of a Normal RV are 1 and 3, respectively.

- Compute the mean

Exercise 5

First and third quartile of a Normal RV are 1 and 3, respectively.

- Compute the mean

The Normal distribution is simmetric, median and mean are the same. The mean is equal to 2

Exercise 5

First and third quartile of a Normal RV are 1 and 3, respectively.

- Compute the variance

Exercise 5

First and third quartile of a Normal RV are 1 and 3, respectively.

- Compute the variance

The percentile $z_{0.75} = \frac{3-2}{\sigma} = \frac{1}{\sigma}$

Exercise 5

First and third quartile of a Normal RV are 1 and 3, respectively.

- Compute the variance

The percentile $z_{0.75} = \frac{3-2}{\sigma} = \frac{1}{\sigma}$

So, we know that

$$P(Z < z_{0.75}) = 0.75$$

By searching into the table:

$$P(Z < z_{0.75}) = 0.75 \rightarrow z_{0.75} = \frac{0.67 + 0.68}{2} = \frac{1}{\sigma}$$

Exercise 5

First and third quartile of a Normal RV are 1 and 3, respectively.

- Compute the variance

$$\sigma^2 = 2.20^2 = 4.84$$

Exercise 5

First and third quartile of a Normal RV are 1 and 3, respectively.

- Compute the variance of $Y = 3X - 2$

Exercise 5

First and third quartile of a Normal RV are 1 and 3, respectively.

- Compute the variance of $Y = 3X - 2$

$$\sigma_Y^2 = 3^2 \times \sigma_X^2 = 9 \times 4.84 = 43.56$$