TaskB1

zerofrom

2024-11-25

B.1

(1) Value of a

Considering chat x follows the probability density function:

$$p_\lambda(x) = \begin{cases} ae^{-\lambda(x-b)} & \text{if} x \geq b, \\ 0 & \text{if} x < b, \end{cases}$$

The probability density function $p_\lambda(x)$ must follows:

$$\int_{-\infty}^{\infty} p_\lambda(x) dx = 1$$

Since $p_\lambda(x) = 0$ when $x < b$, it follows that:

$$\int_{b}^{\infty} p_\lambda(x) dx = 1$$

Replacing expression (1) into expression (3) yields:

$$\int_{b}^{\infty} a\, e^{-\lambda(x-b)} dx = 1$$

Assuming that $u = x - b$, it can be obtained that:

$$x = u + b, dx = du$$

Assuming that $x = b$, then $u = 0$, therefore the upper and lower limits of integration become $0 \to \infty$, substituting expression (s) :

$$\int_{0}^{\infty} a\, e^{-\lambda u} du = 1$$

Using the exponential integration formula:

$$\int_{0}^{\infty} e^{-au}\, du = \frac{1}{a}$$

Thus:

$$a \cdot \frac{1}{\lambda} = 1$$

The solution is:

$$a = \frac{1}{\lambda}$$

In conclusion, the value of a is $\frac{1}{\lambda}$.

(2) i) Population Mean $E[X]$:

From the definition of the mean, there is:

$$E[x] = \int_{-\infty}^{\infty} x \cdot p_\lambda(x)dx$$

Replacing $p_\lambda(x)$ and $a = \lambda$:

$$E[x] = \int_{b}^{\infty} x \cdot \lambda e^{-\lambda(x-b)}dx$$

Assuming $u = x - b$, then $x = u + b$, and $dx = du$:

$$\begin{aligned} E[X] &= \int_{0}^{\infty} (u + b) \cdot \lambda e^{-\lambda u}\, du \\ &= \lambda \cdot \int_{0}^{\infty} u\, e^{-\lambda u}\, du + \lambda b \int_{0}^{\infty} e^{-\lambda u}\, du \end{aligned}$$

Integral formula for exponential functions:

$$\int_{0}^{\infty} e^{-ax}\, dx = \frac{1}{a}, \int_{0}^{\infty} x\, e^{-ax} dx = \frac{1}{a^2}$$

Replacing expression (t) into expression (3), there is:

$$\begin{aligned} E[x] &= \lambda \cdot \frac{1}{\lambda^2} + \lambda b \cdot \frac{1}{\lambda} \\ &= \frac{1}{\lambda} + b \end{aligned}$$

In conclusion, the population mean $E[x] = \frac{1}{\lambda} + b$.

ii) Standard Deviation σx :

The standard deviation formula:

$$\sigma x = \sqrt{\mathrm{Var}(x)} = \sqrt{E[x^2] - (E[x])^2}$$

From the derivation of $E[x]$, it follows that:

$$E[x] = \frac{1}{\lambda} + b$$

The formula for $E[x^2]$ is:

$$E[x^2] = \int_b^\infty x^2 \cdot p_\lambda(x)dx$$

Substituting $P_\lambda(x)$:

$$E[x^2] = \int_b^\infty x^2 \cdot \lambda e^{-\lambda(x-b)}dx$$

Assuming $u = x - b$. then $x = u + b$, and $dx = du$:

$$\begin{aligned} E[x^2] &= \int_0^\infty (u + b)^2 \cdot \lambda e^{-\lambda u}du \\ &= \lambda \int_0^\infty (u^2 + 2bu + b^2) \cdot e^{-\lambda u}du \\ &= \lambda \int_0^\infty u^2\, e^{-\lambda u}du + 2\lambda b \int_0^\infty u\, e^{-\lambda u}du + \lambda b^2 \int_0^\infty e^{-\lambda u}\, du \end{aligned}$$

Integrals formula for exponential functions:

$$\begin{aligned} \int_0^\infty e^{-ax}\, dx &= \frac{1}{a} \\ \int_0^\infty x\, e^{-ax}dx &= \frac{1}{a^2} \\ \int_0^\infty x^2\, e^{-ax}dx &= \frac{2}{a^3} \end{aligned}$$

Replacing formulas (6) into expression (5), there is:

$$\begin{aligned} E[x^2] &= \lambda \cdot \frac{2}{\lambda^3} + 2\lambda b \cdot \frac{1}{\lambda^2} + \lambda b^2 \cdot \frac{1}{\lambda} \\ &= \frac{2}{\lambda^2} + \frac{2b}{\lambda} + b^2 \end{aligned}$$

Replacing expression (2) and (7) into expression (1):

$$
\begin{aligned}
\sigma_x &= \sqrt{\left(\frac{2}{\lambda^2} + \frac{2b}{\lambda} + b^2\right) - \left(\frac{1}{\lambda} + b\right)^2} \\
&= \sqrt{\left(\frac{2}{\lambda^2} + \frac{2b}{\lambda} + b^2\right) - \left(\frac{1}{\lambda^2} + \frac{2b}{\lambda} + b^2\right)} \\
&= \frac{1}{\lambda}
\end{aligned}
$$

In conclusion, the standard deviation of X is $\sigma_x = \frac{1}{\lambda}$.

(3) i) Cumulative distribution function (CDF): Considering the definition of CDF, it follows that:

$$F_x(x) = P(x \leq x) = \int_{-\infty}^{\infty} p_x(t)dt.$$

If $x < b$, then $p_x(t) = 0$, thus there is:

$$F_X(x) = \int_{-\infty}^{x} 0\,dt = 0$$

If $x \geq b$, then $p_x(t) = \lambda e^{-\lambda(t-b)}$, thus there is:

$$F_X(x) = \int_{b}^{x} \lambda\, e^{-\lambda(t-b)}dt$$

Assuming that $a = t - b$, then $t = u + b$. and $dt = du$, the upper and Cower limit of integration becomes $0 \rightarrow x - b$.

$$
\begin{aligned}
F_X(x) &= \int_{0}^{x-b} \lambda\, e^{-\lambda u}du \\
&= \lambda \cdot \int_{0}^{x-b} e^{-\lambda u}\,du \\
&= \lambda\left[-\frac{1}{\lambda}e^{-\lambda u}\right]_{0}^{x-b} \\
&= \lambda \cdot \left(-\frac{1}{\lambda}e^{-\lambda(x-b)} + \frac{1}{\lambda}e^{0}\right) \\
&= \lambda \cdot \left(-\frac{1}{\lambda}e^{-\lambda(x-b)} + \frac{1}{\lambda}\right) \\
&=- e^{-\lambda(x-b)} + 1
\end{aligned}
$$

In conclusion, the cumulative distribution function for the random variable x is

$F_x(x) = \begin{cases} \mathbf{0}, & \text{if} x < b \\ \mathbf{1} - e^{-x(x-b)}, & \text{if} x \geq b \end{cases}$

ii) Quartile Function (QF)

Considering the CDF of random variable X :

$$F_X(x) = \begin{cases} 0, \text{if} x < b \\ 1 - e^{-\lambda(x-b)}, \text{if} x \geq b \end{cases}$$

Quartile Function $Q(p)$ satisfies:

$$F_X\big(Q(p)\big) = p, \text{if} 0 < p < 1$$

Replacing (1) into cl), it follows that:

$$1 - e^{-\lambda(x-b)} = p$$

Thus there is:

$$e^{-\lambda(x-b)} = 1 - p$$

Taking the natural logarithm of both sides:

$$-\lambda(x - b) = \ln(1 - p)$$

Thus

$$x = b - \frac{\ln(1-p)}{\lambda}$$

In conclusion, the Quartile Function of random variable X is $Q(p) = b - \frac{\ln(1-p)}{\lambda}, 0 < p < 1$

(4) Maxmum Likehood estimat :

The definition of likehood function is:

$$\begin{aligned} L(\lambda) &= \prod_{i=1}^{n} p_\lambda(x_i) \\ &= \prod_{i=1}^{n} \lambda e^{-\lambda(x_i-b)} \\ &= \lambda^n \prod_{i=1}^{n} e^{-\lambda(x_i-b)} \\ &= \lambda^n e^{-\lambda \sum_{i=1}^{n}(x_i-b)} \end{aligned}$$

Taking the natural logarithm of both sides:

$$\ln L(\lambda) = n\ln\lambda - \lambda \sum_{i=1}^{n} (x_i - b)$$

Derivativing with respect to λ :

$$\frac{d\ln L(\lambda)}{d\lambda} = \frac{n}{\lambda} - \sum_{i=1}^{n} (x_i - b)$$

Assuming that the derivative is zero, it follows that:

$$\frac{n}{\lambda} - \sum_{i=1}^{n} (x_i - b) = 0$$

Calculating the value of λ is:

$$\lambda = \frac{n}{\sum_{i=1}^{n} (x_{i-b})}$$

In conclusion, the maxmum likehood estimate for x is $\lambda_{\text{MLE}} = \frac{n}{\sum_{i=1}^{n} (x_i - b)}$

(5) Given the sample, compute and display the maximum likelihood estimate λMLE of the parameter λ.

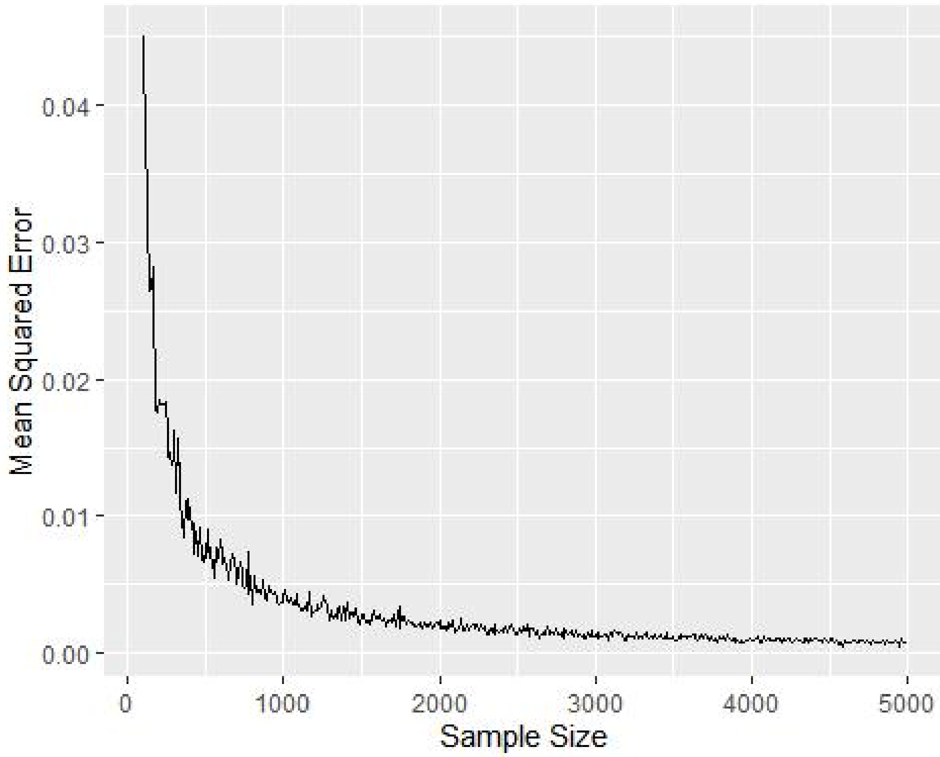
```
## [1] "The maximum likelihood estimate lambda_MLE is:  0.0198842607985
72"
```

(6) Compute the Bootstrap confidence interval

```
## [1] "The Bootstrap confidence interval is: [ 0.0191115894121709 , 0.
0206792527343937 ]\n"
```

(7)) Conduct a simulation study to explore the behaviour of the maximum likelihood estimator λMLE for λ on simulated data X1, · · · , Xn

```
##   SampleSize        MSE
## 1        100 0.04507438
## 2        110 0.03664986
## 3        120 0.03907616
## 4        130 0.03197187
## 5        140 0.02650689
```



B2

(1) The formula for the probability mass function

The range of possible values of the discrete ranom variable x is:

$$x \in \{-2,0,2\}$$

Event 1: Draw 2 blue balls ($X=-2$)

$$C(b,2) = \frac{b(b-1)}{2}, C(a+b,2) = \frac{(a+b)(a+b-1)}{2}$$
$$\therefore P(x=-2) = \frac{C(b,2)}{C(a+b,2)} = \frac{b(b-1)}{(a+b)(a+b-1)}$$

Event 2: Draw 1 red ball and 1 blue ball ($x = 0$)

$$C(a,1) \cdot C(b,1) + C(b,1) \cdot C(a,1) = 2 \cdot a \cdot b$$
$$\therefore \quad P(X=0) = \frac{C(a,1)C(b,1) + C(b,1)C(a,1)}{C(a+b,2)}$$
$$= \frac{2ab}{(a+b)(a+b-1)}$$

Event 3: Draw 2 red balls ($x = 2$)

$$\begin{aligned} C(a,2) &= \frac{a(a-1)}{2} \\ \therefore P(x=2) &= \frac{C(a,2)}{C(a+b,2)} = \frac{a(a-1)}{(a+b)(a+b-1)} \end{aligned}$$

In conclustion, the formula for the probability mass function pros:

$$P(x) = \begin{cases} \dfrac{b(b-1)}{(a+b)(a+b-1)} & ,x=-2 \\ \dfrac{2ab}{(a+b)(a+b-1)} & ,x=0 \\ \dfrac{a(a-1)}{(a+b)(a+b-1)} & ,x=2 \end{cases}$$

(2) The expression of the expectation

According to the expectation formula for random variables:

$$E(X) = \sum_{x\in\{-2,0,2\}} x \cdot P(X=x)$$

Replacing $P(x)$ into (1) , there is:

$$\begin{aligned} E(x) &= -2 \cdot \frac{b(b-1)}{(a+b)(a+b-1)} + 2 \cdot \frac{a(a-1)}{(a+b)(a+b-1)} \\ &= \frac{2a(a-1) - 2b(b-1)}{(a+b)(a+b-1)} \\ &= \frac{2(a^2 - b^2 - a + b)}{(a+b)(a+b-1)} \end{aligned}$$

In conclusion, the expectation $E(X)$ of X is:

$$E(X) = \frac{2(a^2 - b^2 - a + b)}{(a+b)(a+b-1)}$$

(3)The expression of the variance Var(X)

According to the formula of $E(X^2)$:

$$E(X^2) = \sum_{x\in\{-2,0,2\}} x^2 \cdot P(X=x)$$

Replacing $P(x)$ in to (1), there is:

$$
\begin{aligned}
E(x^2) &= 4 \cdot \frac{b(b-1)}{(a+b)(a+b-1)} + 4 \cdot \frac{a(a-1)}{(a+b)(a+b-1)} \\
&= \frac{4(a^2+b^2-a-b)}{(a+b)(a+b-1)}
\end{aligned}
$$

According to the formula of Var(X) :

$$\mathrm{Var}(x) = E(x^2) - [E(x)]^2$$

Replacing $E(x^2)$ and $E(x)$ into (3), it follows that:

$$
\begin{aligned}
\mathrm{Var}(x) &= \frac{4(a^2+b^2-a-b)}{(a+b)(a+b-1)} - \left[\frac{2(a^2-b^2-a+b)}{(a+b)(a+b-1)}\right]^2 \\
&= \frac{4(a^2+b^2-a-b)(a+b)(a+b-1) - 4(a^2-b^2-a+b)^2}{(a+b)^2(a+b-1)^2}
\end{aligned}
$$

In conclusion, the expression of the variance Var(X) is:

$$\mathrm{Var}(x) = \frac{4(a^2+b^2-a-b)(a+b)(a+b-1) - 4(a^2-b^2-a+b)^2}{(a+b)^2(a+b-1)^2}$$

(4)Write a function called compute_expectation_X that takes a and b as inputs and outputs the expectation E(X). Write a function called compute_variance_X that takes a and b as input and outputs the variance Var(X).

(5) The expression of the expectation of the random variable X

According to the linear nature of the mean of a random variable:

$$E(\bar{X}) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E\,(X_i)$$

$\because$ X_i are i.i.d., thus $E(X_i) = E(x)$, thus:

$$E(\bar{X}) = \frac{1}{n} \cdot n \cdot E(X) = E(X)$$

In conclusion, the expression of the expectation of the random variable $\bar{x}$ is: $E(\bar{X}) = \frac{2(a^2-b^2-a+b)}{(a+b)(a+b-1)}$

(6) The expression of the variance of the random variable X

Considering the effect of a linear transformation of a random variable on the variance :

$$ \begin{equation} \operatorname{Var}(a X+b)=a^{2} \operatorname{Var}(X) \tag{1}\\ \end{equation} $$

thus:

$$
\begin{aligned}
\operatorname{Var}(\bar{X}) &= \operatorname{Var}\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \\
&= \frac{1}{n^2}\operatorname{Var}\left(\sum_{i=1}^{n} X_i\right) \\
&= \frac{1}{n^2}\cdot n\operatorname{Var}(X) \\
&= \frac{1}{n}\operatorname{Var}(X)
\end{aligned}
$$

In conclusion, the expression of the variance of the random variable $\bar{X}$ is

$$\operatorname{Var}(\bar{X}) = \frac{4(a^2+b^2-a-b)(a+b)(a+b-1)-4(a^2-b^2-a+b)^2}{n(a+b)^2(a+b-1)^2}$$

(7) Create a function called sample_Xs

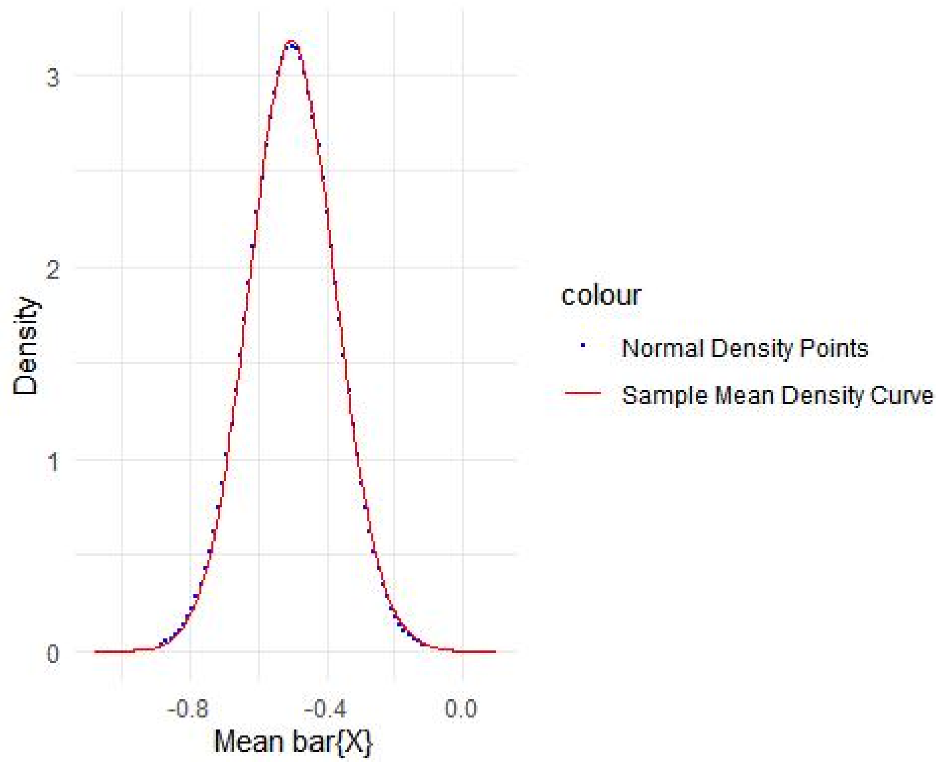
(8) Calculate E(X),Var(X),E({X}),Var({X})

```
## [1] "Expectation E(X): -0.5 \n"

## [1] "Expectation of samples E(bar{X}): -0.501 \n"

## [1] "Variance Var(X): 1.60714285714286 \n"

## [1] "Variance of samples Var(bar{X}): 1.60309503095031 \n"
```

(9) Compute the corresponding sample mean X based on X1, · · · , Xn

(10) Create a scatter plot of the points

```
## Warning: Using `size` aesthetic for lines was deprecated in ggplot2
3.4.0.
## ℹ Please use `linewidth` instead.
## This warning is displayed once every 8 hours.
## Call `lifecycle::last_lifecycle_warnings()` to see where this warnin
g was
## generated.
```



(11) Describe the relationship between the density of X and the function fμ,σ displayed in your plot. Try to explain the reason.

(i) The blue points in the figure depicts the density distribution of the mean of the discrete random variable X over the interval [μ - 3σ, μ + 3σ], satisfying the normal distribution.

(ii) The red curve in the figure is the kernel density estimate obtained by representing the kernel density of the sample mean X within simulation study with 50000 trials.

(iii) It can be seen that the two are very close to each other, proving that the sampling distribution of the sample means tends to the standard normal distribution when the sample size is large enough.