Introduction to Probability & Statistics

Assignment 5, 2024/25

Assignment Questions – answers to be uploaded

AQ1. Let $X_1, \dots, X_{16}$ be an i.i.d. sample from a $\mathsf{N}(3, 1)$ distribution, and let $S = X_1 + X_2 + \cdots + X_{16}$. Express $\mathbb{P}(S < 52)$ in terms of the distribution function Φ of the standard normal distribution.

> Answer
>
> We know that $\mathbb{E}[S] = 16\mathbb{E}[X] = 48$, and $\mathsf{Var}(S) = 16\mathsf{Var}(X) = 16$ (because the X_i are independent, and hence uncorrelated).
> Furthermore, we know that the sum of independent normal distributions has a normal distribution. Thus
>
> $$S \sim \mathsf{N}(48, 16).$$
>
> **[3]**
> We can normalize to obtain a standard normal random variable by subtracting the mean and dividing by the standard deviation. Thus
>
> $$\mathbb{P}(S < 52) = \mathbb{P}\left(\frac{S - 48}{\sqrt{16}} < \frac{52 - 48}{\sqrt{16}}\right) = \Phi(1).$$
>
> **[2]**

AQ2. Let $Y_1, Y_2, \dots$ be an i.i.d. sequence of random variables, each with a $\mathsf{Uniform}(0, 3)$ distribution. Define a new sequence of random variables $X_1, X_2, \dots$ by

$$X_n = \frac{1}{n}\sum_{i=1}^{n} Y_i^2.$$

Using the Law of Large Numbers, determine the value of $a \in \mathbb{R}$ for which $\mathbb{P}(\lim_{n\to\infty} X_n = a) = 1$.

> Answer
>
> The random variable X_n is just the sample mean of the random variables $Y_1^2, \dots, Y_n^2$. These are i.i.d. and clearly have finite mean and variance (since Y_i^2 can only take values in the finite set $[0, 9]$).
> The (strong) law of large numbers says that, with probability one, X_n will converge to $\mathbb{E}[Y^2]$. **[3]**
> Finally, we calculate
>
> $$\mathbb{E}\left[Y^2\right] = \int_0^3 y^2 \frac{1}{3} dy = 3,$$
>
> and so the required answer is $a = 3$. **[2]**

AQ3. Suppose the random variables X_1, X_2 and X_3 all have the same expectation μ. For what values of a and b is

$$M = -4(X_1 - 2) + 9(X_2 - 1) + aX_3 + b$$

an unbiased estimator for μ?

Answer

M is an unbiased estimator for μ if $\mathbb{E}[M] = \mu$ for any value of μ. **[1]** We find

$$\begin{aligned}
\mathbb{E}[M] &= \mathbb{E}[-4(X_1-2)+9(X_2-1)+aX_3+b] \\
&= -4(\mathbb{E}[X_1]-2)+9(\mathbb{E}[X_2]-1)+a\mathbb{E}[X_3]+b \\
&= -4(\mu-2)+9(\mu-1)+a\mu+b \\
&= (5+a)\mu+(b-1).
\end{aligned}$$

[2]
Thus $\mathbb{E}[\mu] = \mu$ if and only if $a = -4$ and $b = 1$. **[2]**

AQ4. From a dataset $x_1, \ldots, x_{10}$ it has been calculated that

$$\sum_{i=1}^{10} x_i = 491, \quad \sum_{i=1}^{10} (x_i - \bar{x}_{10})^2 = 41.$$

You model the dataset as a random sample from the normal distribution with mean μ and variance σ^2.

a) Assume that both μ and σ^2 are unknown. Determine a 95% confidence interval for the mean μ. You can use that $t_{9,0.025} \approx 2.26$.

b) Now assume that it is known that the variance is $\sigma^2 = 5$. Give a 95% confidence interval for the mean μ in this case. You can use that $z_{0.025} \approx 1.96$.

Answer

a) Since the variance is unknown, we use the interval

$$\left(\bar{x}_n - t_{n-1,\alpha/2}\frac{s_n}{\sqrt{n}}, \bar{x}_n + t_{n-1,\alpha/2}\frac{s_n}{\sqrt{n}}\right)$$

We first calculate $\bar{x}_n = 491/10 = 49.1$ and

$$s_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x}_n)^2 = \frac{1}{10-1}41 = \frac{41}{9}.$$

We're told that $t_{n-1,\alpha/2} = t_{9,0.025} = 2.26$ and

$$\begin{aligned}
\left(\bar{x}_n - t_{n-1,\alpha/2}\frac{s_n}{\sqrt{n}}, \bar{x}_n + t_{n-1,\alpha/2}\frac{s_n}{\sqrt{n}}\right) &\approx \left(49.1 - 2.26\frac{\sqrt{41}}{3\sqrt{10}}, 49.1 + 2.26\frac{\sqrt{41}}{3\sqrt{10}}\right) \\
&\approx (47.57, 50.63).
\end{aligned}$$

[3]

b) When σ^2 is known, we use the confidence interval

$$\begin{aligned}
\left(\bar{x}_n - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{x}_n + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) &\approx \left(49.1 - 1.96\frac{\sqrt{5}}{\sqrt{10}}, 49.1 + 1.96\frac{\sqrt{5}}{\sqrt{10}}\right) \\
&\approx (47.71, 50.49).
\end{aligned}$$

[2]