MATH7501: Exercise 8 Solutions

Dinesh Kalamegam

March 25, 2018

1 Question 1 (6 MARKS)

1.1 Showing s^2 is a unbiased estimator of σ^2 also states its variance

Given $s^2 \sim \Gamma(\alpha, \lambda)$ for $\alpha = \frac{n-1}{2}$ and $\lambda = \frac{n-1}{2\sigma^2}$

$$\begin{aligned}
E(s^2) &= \frac{\alpha}{\lambda} \\
&= \frac{(n-1)/2}{(n-1)/2\sigma^2} \\
&= \frac{2\sigma^2}{2} \\
&= \boxed{\sigma^2}
\end{aligned}$$

Then the $Bias(s^2) = E(s^2) - \sigma^2 = 0$ means that s^2 is an unbiased estimator for σ^2

$$\begin{aligned}
Var(s^2) &= \frac{\alpha}{\lambda^2} \\
&= \frac{(n-1)/2}{(n-1)^2)/4\sigma^2} \\
&= \boxed{\frac{2\sigma^4}{n-1}}
\end{aligned}$$

1.2

1.2.1 Find expressions for the bias and variance of T_k as an estimator of σ^2

Consider $T_k = ks^2$.

$$E(T_k) = E(ks^2) = kE(s^2) = k\sigma^2$$

$$Var(T_k) = Var(ks^2) = k^2 Var(s^2) = \boxed{\frac{2k^2\sigma^4}{n-1}}$$

$Bias(T_k) = E(T_k) - \sigma^2$ as an estimator of σ^2 which then gives us $k\sigma^2 - \sigma^2 = \boxed{\sigma^2(k-1)}$

Then the Mean Square Error (MSE) of T_k is

$$\begin{aligned} MSE(T_k) &= Bias^2(T_k) + Var(T_k) \\ &= (k-1)^2\sigma^2 + \frac{2k^2\sigma^4}{n-1} \end{aligned}$$

1.3 Minimise the error

To minimise the MSE in terms of k consider

$$\frac{d}{dk} MSE(T_k) = 2\sigma^4(k-1) + \frac{4\sigma^4}{n-1}k$$

Then by setting this derivate to zero and assuming $\sigma \neq 0$ we obtain

$$\begin{aligned} 2\sigma^4(k-1) + 4\sigma^4 k = 0 \\ &= 2\sigma^4\left((k-1) + \frac{2k}{n-1}\right) = 0 \\ &\implies \boxed{k = \frac{n-1}{n+1}} \end{aligned}$$

To confirm that this leads to a minimum value of the MSE consider the second derivative

$$\frac{\partial^2}{\partial k} MSE(T_k) = 2\sigma^4 + \frac{4\sigma^4}{n-1} > 0$$

So we have a minimum

2 Question 2 (4 MARKS)

2.1

$Z \sim N(0,1)$ and $U \sim \chi^2_\nu$ So we have that $Z^2 \sim \chi^2_1$. Then by defintion

$$T = \frac{Z}{\sqrt{\frac{U}{\nu}}}$$

$$\boxed{\text{Is a } t_\nu \text{ distribution with } \nu \text{ degrees of freedom}}$$

Then for the next Part

$$T^2 \sim \frac{Z^2}{U/\nu} = \frac{Z^2/1}{U/\nu}$$

$$= \frac{\chi_1^2}{\chi_\nu^2}$$

$$= \boxed{\text{distributed by } F_{1,\nu}}$$

2.2

2.2.1 $P(Y > 4)$

Given that $Y \sim F_{1,5}$ we can use 2.1 to give $Y = T^2$ and $T \sim t_5$.

The required probability is then

$$\begin{aligned}
P(Y > 4) &= P(T^2 > 4) \\
&= P(T > 2) + P(T < -2) \\
&= 2P(T > 2) \text{ by symmetry of the } t \text{ distribution} \\
&= 2(1 - P(T < 2)) \\
&= 2(1 - 0.9490) \\
&= \boxed{0.102}
\end{aligned}$$

2.2.2 find c s.t. $P(Y > c) = 0.01$

This follows the steps above and this leads us to

$$\begin{aligned}
P(Y > c) &= P(T^2 > c) \\
&= P(T > \sqrt{c}) + P(T < -sqrtc) \\
&= 2P(T > \sqrt{c}) \text{ by symmetry of the } t \text{ distribution} \\
&= 2(1 - P(T < \sqrt{c})) = 0.01 \\
&\implies 1 - P(T < \sqrt{c}) = 0.005 = 0.5\% \\
&\text{Recall that } T \sim t_5 \text{ so from table 10 read from row } \nu = 5 \text{ and column } 0.5 \\
&\implies \sqrt{c} = 4.032 \\
&\implies \boxed{c = 16.257 \text{ to 3sf}}
\end{aligned}$$

3 Question 3 (10 MARKS)

3.1

3.1.1 Stem and Leaf diagram

```
Key 197| 8 = 197.8 cm

  197 | 8  9  9
  198 | 0  2  3
  199 | 2  6  7
  200 | 2  2  5
  201 | 4  8
  202 | 0
  203 | 3
```

3.1.2 Sample mean and variance

For this data we have $n = 16$ and

$$\sum_{i=1}^{16} X_i = 3196$$
$$\sum_{i=1}^{n} X_i^2 = 638445.1$$

Then the sample mean

$$\begin{aligned}\bar{X} &= \frac{3196}{16} \\ &= \boxed{197.75}\end{aligned}$$

And the sample variance

$$\begin{aligned}s^2 &= \frac{\sum_{i=1}^{16} X_i^2 - n\bar{X}}{n-1} \\ &= \frac{1}{15}(638445.1 - 16(199.75^2)) \\ &= \boxed{2.94}\end{aligned}$$

3.2 95% confidence interval

From the notes we have that $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$ i.e $Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$

At a 95% confidence interval for Z is $(-1.96, 1.96)$ and from the tables we have that

$P(Z < 1.96) = 0.975 \wedge P(Z > 1.96) = 0.025$

Then the confidence interval for μ can be found using $-1.96 < \frac{\bar{X}-\mu}{\sigma/\sqrt{n}} < 1.96$ we also have $\bar{X}$ as well as the fact that question gives that $\sigma = 1$

$$\frac{\bar{X} - 1.96}{\sigma/\sqrt{n}} < \mu < \frac{\bar{X} + 1.96}{\sigma/\sqrt{n}}$$

$$\frac{199.75 - 1.96}{1/4} < \mu < \frac{199.75 + 1.96}{1/4}$$

$$199.26 < \mu < 200.24$$

This gives us the confidence interval $\boxed{(199.26, 200.24)\text{cm}}$ The interval includes the nominal value of 200 cm and hence data is consistent with underlying mean of 200cm