(i) Sample median is not affected by the fact that the last two observations are censored.

It is therefore given by the 5.5\textsuperscript{th} ranked observation, i.e. \((355 + 379) / 2 = 367\) days. \[2\]

(ii) We know that the last two observations have minimum values 432 and 463.

Using these two values the sample mean would be equal to \(3679/10 = 367.9\).

So, the sample mean is at least equal to 367.9 days. \[2\]

[Total 4]

2

(i) \[M_X(t) = E(e^{tX}) = \frac{1}{2} \int_{-\infty}^{0} e^{(t+1)x} dx + \frac{1}{2} \int_{0}^{\infty} e^{(t-1)x} dx\]

\[= \frac{1}{2} \left[ \frac{e^{(t+1)x}}{t+1} \right]_{-\infty}^{0} + \frac{1}{2} \left[ \frac{e^{(t-1)x}}{t-1} \right]_{0}^{\infty}\]

and for \(|t|<1\)

\[M_X(t) = \frac{1}{2} \left( \frac{1}{t+1} - \frac{1}{t-1} \right) = \frac{1}{1-t^2}\] \[3\]

(ii) \[M_X'(t) = \left( (1-t^2)^{-1} \right)' = -(1-t^2)^{-2}(-2t) = 2t(1-t^2)^{-2}\]

\[\Rightarrow E(X) = M_X'(0) = 0\]

\[M_X''(t) = \left( 2t(1-t^2)^{-2} \right)' = 2(1-t^2)^{-2} + 2t(-2)(1-t^2)^{-3}(-2t)\]

\[= 2(1-t^2)^{-2} + 8t^2(1-t^2)^{-3}\]

\[\Rightarrow E(X^2) = M_X''(0) = 2\]

\[V(X) = E(X^2) - E^2(X) = 2\]

(Alternatively, based on a series expansion:

\[M_X(t) = 1 + t^2 + t^4 + \ldots \Rightarrow E(X) = 0\] and \(E(X^2) = 2\) and the variance follows.) \[3\]

[Total 6]
3

(i) \( P(Y = 2) = 0.25 + 0.05 = 0.3 \) \[1\]

(ii) \( P(X = 0) = 0.75 \) and
\[
P( \{X = 0\} \cap \{Y = 2\} ) = 0.25 \neq 0.225 = 0.3 \times 0.75
\]

Therefore \( X \) and \( Y \) are not independent.

(Other joint probabilities could be used, but not those with \( Y = 1 \).) \[2\]

(iii) The probability function is

\[
P( R = r ) \quad 0.25 \quad 0.05 \quad 0.3 \quad 0.1 \quad 0.2 \quad 0.1
\]

[Total 6]

4

(i) The random variables \( X_1, \ldots, X_n \) are independent
and identically distributed with \( X_i \sim N(\mu, \sigma^2) \) \[2\]

(ii) \( \bar{X} \) and \( S^2 \) are independent
\[
\bar{X} \sim N(\mu, \sigma^2 / n)
\]
\[
\frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}
\]

(iii) \( t_k = N(0,1) / \sqrt{\frac{\chi^2_k}{k}} \) where \( N(0,1) \) and \( \chi^2_k \) are independent

This result can be applied here, and we get \( \frac{\bar{X} - \mu}{S / \sqrt{n}} \sim t_{n-1} \) \[2\]

[Total 7]

5

(i) \( P(\mu > 180) = P(\mathcal{N}(187, 10^2) > 180) \)
\[
= P\left( N(0,1) > \frac{180 - 187}{10} \right)
\]
= P(N(0,1) > -0.7)
= 0.75804 \[2\]

(ii) We know that \(\mu|\bar{x} \sim N(\mu_*, \sigma_2^2)\)

Where \(\mu_* = \left(\frac{80 \times 182 + 187}{15^2 + 10^2}\right) = 182.14\)

And \(\sigma_2^2 = \frac{1}{15^2 + 10^2} = 2.73556 = 1.6540^2\)

so \(P(\mu > 180) = P(182.14, 1.654^2 > 180)\)

\[= P\left(N(0,1) > \frac{180 - 182.14}{1.654}\right)\]

\[= P(N(0,1) > -1.29192)\]

\[= 0.38 \times 0.9032 + 0.62 \times 0.90147\]

\[= 0.90180 \[4\]\]

(iii) The probability has risen, reflecting our much greater certainty over the value of \(\mu\) as a result of taking a large sample.

This is despite the fact that our mean belief about \(\mu\) has fallen, which a priori might make a lower value of \(\mu\) more likely.

The posterior distribution has thinner tails / lower volatility, since we have increased credibility around the mean. \[2\]

[Total 8]

6

Let the prior distribution of \(\mu\) have a Gamma distribution with parameters \(\alpha\) and \(\lambda\) as per the tables.

Then \(\frac{\alpha}{\lambda} = 50\) and \(\frac{\alpha}{\lambda^2} = 15^2\)

Then dividing the first by the second \(\lambda = \frac{50}{15^2} = 0.22222\)

And so \(\alpha = 50 \times 0.22222 = 11.11111\)
The posterior distribution of $\mu$ is then given by

\[
f(\mu|x) \propto f(x|\mu) f(\mu)
\]

\[
\propto e^{-10\mu} \times \mu^{-630} \times e^{10.1111} \times e^{-0.22222\mu}
\]

\[
\propto e^{640.11111} \times e^{-10.22222\mu}
\]

Which is the pdf of a Gamma distribution with parameters $\alpha' = 641.11111$ and $\lambda' = 10.22222$

Now under all or nothing loss, the Bayesian estimate is given by the mode of the posterior distribution. So we must find the maximum of

\[
f(x) = x^{640.11111} e^{-10.22222x}
\]

(we may ignore constants here)

Differentiating:

\[
f'(x) = e^{-10.22222x} \left( -10.22222x^{640.11111} + 640.11111x^{639.11111} \right)
\]

\[
= x^{639.11111} e^{-10.22222x} (-10.22222x + 640.11111)
\]

And setting this equal to zero we get

\[
x = \frac{640.11111}{10.22222} = 62.62
\]

7

(i) The link function here is $g(\mu) = \log \mu$.

(ii) (a) The linear predictor is $\alpha_i + \beta_i x$ where the intercept $\alpha_i$ for $i = 1, 2$ depends on gender.

(b) The linear predictor is $\alpha_i + \beta_i x$ where $\beta_i$ for $i = 1, 2$ also depends on gender, so that both parameters depend on gender.

[Total 5]

8

(i) From the definition of the gamma density given in the question

\[
f(y) = \frac{\alpha^\alpha}{\mu^\alpha \Gamma(\alpha)} y^{\alpha-1} e^{-\frac{\alpha}{\mu} y}
\]
\[= \exp \left[ \left( -\frac{y}{\mu} - \log \mu \right) \alpha + (\alpha - 1) \log y + \alpha \log \alpha - \log \Gamma(\alpha) \right] \]

\[= \exp \left[ \frac{(y_0 - b(\theta))}{a(\varphi)} + c(y, \varphi) \right] \]

where:
\[\theta = -\frac{1}{\mu}\]
\[\varphi = \alpha\]
\[a(\varphi) = \frac{1}{\varphi}\]
\[b(\theta) = -\log (-\theta)\]
\[c(y, \varphi) = (\varphi - 1) \log y + \varphi \log \varphi - \log \Gamma(\varphi).\]

Hence the distribution has the right form for a member of an exponential family.

The natural parameter is \(-\frac{1}{\mu}\). The canonical link function is \(\frac{1}{\mu}\). [5]

(ii) Using the information given, we can calculate the deviance differences and compare that with the differences of the degrees of freedom for each of the nested models. If the decrease in the deviances is greater than twice the difference in degrees of freedom this suggests an improvement.

<table>
<thead>
<tr>
<th>Model</th>
<th>Scaled Deviance</th>
<th>Degrees of freedom</th>
<th>Difference in scaled deviance</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>900</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>789</td>
<td>10</td>
<td>111</td>
</tr>
<tr>
<td>Age +location</td>
<td>544</td>
<td>7</td>
<td>245</td>
</tr>
<tr>
<td>Age * location</td>
<td>541</td>
<td>1</td>
<td>3</td>
</tr>
</tbody>
</table>

From the table we can see that the interaction model does not indicate any improvement hence the recommended model would be Age +location. [6]

[Total 11]
In thousands:

(i) \[
\left[ 21 - t_{0.025,24} \frac{2.5}{5}, \ 21 + t_{0.025,24} \frac{2.5}{5} \right] = \left[ 21 - 2.064 \frac{1}{2}, \ 21 + 2.064 \frac{1}{2} \right]
[19.968, \ 22.032]
\]

(ii) \( H_0 : \alpha \leq 20 \) vs \( H_1 : \alpha > 20 \)

(or, \( H_0 : \alpha = 20 \) vs \( H_1 : \alpha > 20 \))

Test statistic:

\[
t = \frac{\bar{x} - \alpha_0}{s / \sqrt{n}} = \frac{21 - 20}{2.5 \times 0.2} = 2 > 1.711 = t_{0.05,24}
\]

We reject the null hypothesis.

(iii) \[
\frac{21 - \alpha_0}{2.5 \times 0.2} = 1.711, \quad 21 - \alpha_0 = 0.8555, \quad \alpha_0 = 20.1445
\]

(iv) Test \( H_0 : \lambda = 0.6 \) vs \( H_1 : \lambda \neq 0.6 \)

Test statistic (based on normal approximation to Poisson) is:

\[
z = \frac{\bar{x} - \lambda_0}{\sqrt{\lambda_0 / n}} = \frac{0.5 - 0.6}{\sqrt{0.6 / 100}} = -0.1 \in [-1.96, 1.96]
\]

\[
\frac{0.5 - 0.6}{\sqrt{0.6 / 100}} = -0.1 \in [-1.96, 1.96]
\]

(or, with continuity correction \( z = \frac{0.5 + 0.5 - 0.6}{\sqrt{0.6 / 100}} = -1.226 \))

The null hypothesis \( H_0 : \lambda = 0.6 \) cannot be rejected for the year 2011.

(v) Test \( H_0 : \lambda_{2012} \leq \lambda_{2011} \) vs \( H_1 : \lambda_{2012} > \lambda_{2011} \)

(or, \( H_0 : \lambda_{2012} = \lambda_{2011} \) vs \( H_1 : \lambda_{2012} > \lambda_{2011} \))
Overall sample mean \( \hat{\lambda} = 0.55 \)

Test statistics now is:

\[
\frac{\hat{\lambda}_{2012} - \hat{\lambda}_{2011}}{\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} = \frac{0.6 - 0.5}{\sqrt{1.1/100}} = \frac{0.1}{0.104} = 0.9535 < 1.64
\]

The null hypothesis \( H_0 : \lambda_{2012} \leq \lambda_{2011} \) cannot be rejected at the 5% level. Therefore, we do not have empirical evidence to suggest that the alternative \( \lambda_{2012} > \lambda_{2011} \) is true. [3]

[Total 14]

10

(i) We have:

\[
E[X] = \int \frac{a^c}{x^{a+1}} x dx = ac \int_c^\infty x^{-a} dx = -\frac{ac}{a-1} \left[ x^{-a+1} \right]_c^\infty
\]

and for \( a > 1 \)

\[
E[X] = -\frac{ac}{a-1} (0 - c^{-a+1}) = \frac{ac}{a-1}.
\] [2]

(ii) \( F_X(x) = \int f_X(t) dt = \int_c^x \frac{ac^a}{t^{a+1}} dt \)

which gives

\[
F_X(x) = -\frac{a}{a-1} \left[ x^{-a} \right]_c^x = -\frac{a}{a-1} (x^{-a} - c^{-a}) = 1 - \left( \frac{c}{x} \right)^a, \quad x \geq c
\]

[OR differentiate \( F_X(x) \) to obtain \( f_X(x) \)] [2]

(iii) The likelihood function is given by:

\[
L(a) = \prod_{i=1}^n f_X(x_i) = \prod_{i=1}^n \frac{ac^a}{x_i^{a+1}} = a^n c^n \prod_{i=1}^n x_i^{-(a+1)}
\]

and
\[ l(a) = n \log(a) + na \log(c) - (a + 1) \sum_{i=1}^{n} \log(x_i) \]

For the MLE:

\[ l'(a) = 0 \Rightarrow \frac{n}{a} + n \log(c) - \sum_{i=1}^{n} \log(x_i) = 0 \]

\[ \Rightarrow \hat{a} = \frac{n}{\sum_{i=1}^{n} \log(x_i) - n \log(c)} \sum_{i=1}^{n} \log \left( \frac{x_i}{c} \right) \]

and for \( c = 2.5 \), \( \hat{a} = \frac{n}{\sum_{i=1}^{n} \log \left( \frac{x_i}{2.5} \right)} \) \[3\]

(iv) For the asymptotic variance we use the Cramer-Rao lower bound:

\[ l''(a) = -\frac{n}{a^2}, \quad \text{and} \quad E \left[ l''(a) \right] = -\frac{n}{a^2} \]

\[ \Rightarrow V[\hat{a}] = -\left\{ E \left[ l''(a) \right] \right\}^{-1} = \frac{a^2}{n}. \]

Hence, asymptotically, \( \hat{a} \sim N(a, a^2/n) \). \[4\]

(v) Size of claim in the following year will be given by \( 1.05X \)

So we want \( P(1.05X > 4) = P \left( X > \frac{4}{1.05} \right) = 1 - F_X \left( \frac{4}{1.05} \right) \)

and using \( F_X \) given in the question

\[ P(1.05X > 4) = \left( \frac{1.05 \times 2.5}{4} \right)^6 = 0.0799. \] \[3\]

[Total 14]
(i) Scatterplot with suitable axes and clearly labelled:

There does not appear to be much of a relationship, perhaps a slight increasing linear relationship but it is weak with quite a bit of scatter. 

(ii) \( n = 16 \)

\[
S_{tt} = 1496 - \frac{136^2}{16} = 340
\]

\[
S_{yy} = 12.531946 - \frac{14.160^2}{16} = 0.000346
\]

\[
S_{ty} = 120.518 - \frac{(136)(14.160)}{16} = 0.158
\]

\[
\hat{\beta} = \frac{S_{ty}}{S_{tt}} = \frac{0.158}{340} = 0.0004647
\]

\[
\hat{\alpha} = \bar{y} - \hat{\beta}\bar{t} = \frac{14.160}{16} - (0.0004647) \frac{136}{16} = 0.88105
\]

Fitted line is \( y = 0.88105 + 0.000465t \)

(iii) (a) s.e. (\( \hat{\beta} \)) = \( \sqrt{\frac{\hat{\sigma}^2}{S_{tt}}} \) where \( \hat{\sigma}^2 = \frac{1}{n-2} (S_{yy} - \frac{S_{ty}^2}{S_{tt}}) \)
\[
\sigma^2 = \frac{1}{14} \left( 0.000346 - \frac{0.158^2}{340} \right) = 0.0000195
\]

\[
\therefore \text{s.e.}(\hat{\beta}) = \sqrt{\frac{0.0000195}{340}} = 0.000239
\]

(b) Null hypothesis of “no linear relationship” is equivalent to \(H_0: b = 0\)

We use \(t = \frac{\hat{\beta}}{\text{s.e.}(\hat{\beta})} \sim t_{14}\) under \(H_0: b = 0\)

Observed \(t = \frac{0.000465}{0.000239} = 1.95\) and \(t_{0.025,14} = 2.145\)

So we must accept \(H_0: \) no linear relationship at the 5% level.

(c) 95% CI is \(0.000465 \pm 2.145 \times 0.000239\)

giving \(0.000465 \pm 0.000513\) or \((-0.000048, 0.000978)\) \[5\]

(iv) (a) Observed \(t = \frac{0.000487}{0.000220} = 2.21\) – this is greater than \(t_{0.025,14} = 2.145\)

So we reject \(H_0: \) no linear relationship at the 5% level.

(b) 95% CI is \(0.000487 \pm 2.145 \times 0.000220\)

giving \(0.000487 \pm 0.000472\) or \((0.000015, 0.000959)\)

The two CIs overlap substantially, so there is no evidence to suggest that the slopes are different.

(c) Although the tests have different conclusions at the 5% level, the 100m observed \(t\) is only just inside the critical value of 2.145 and the 200m one is just outside. This in fact agrees with, rather than contradicts, the conclusion that the slopes are not different. \[6\]

[Total 18]

END OF MARKING SCHEDULE