ST1051 Summer Exam 2020

Jack O'Connor 119319446

April 30, 2020

Question 1 (a)

(a) Defective CDs
D = Event chosen CD is defective.
A = Event chosen CD was made in Machine I.

The events A and D are independent iff $P(D) = P(D \mid A)$.

$P(D \mid A)$ is equal to the probability of getting a defective CD if given a CD from Machine 1.

$$P(D) = \frac{15}{100} = 0.15$$
$$P(D \mid A) = \frac{9}{60} = 0.15$$

Since $P(D) = P(D \mid A) = 0.15$ we can conclude that the events D and A are independent.

(b) University Advertisements

Rate λ is per hour therefore time t = 3 minutes converted to 0.05 hours.
$X \sim Pois(\lambda = 50,\, t = 0.05)$

$$P(X = k) = \frac{(\lambda t)^k e^{-\lambda t}}{k!}$$

$$\begin{aligned} P(X = 0) &= \frac{(50 \cdot 0.05)^0 e^{-50 \cdot 0.05}}{0!} \\ &= e^{-2.5} \\ &= 0.08208 \text{ or } 8.208\% \end{aligned}$$

(c) Striking Oil

$X \sim NegBinom(r = 3,\, p = 0.2)$

$$P(X = n) = \binom{n-1}{r-1} p^r (1-p)^{n-r}$$

$$\begin{aligned} P(X = 7) &= \binom{6}{2} (0.2)^3 (0.8)^4 \\ &= 0.04915 \text{ or } 4.915\% \end{aligned}$$

(d) Find The Mean

$\mathbb{E}[X^k] = \lim_{t \to 0} \frac{d^k \phi_X(t)}{dt^k}$

Mean $= \mathbb{E}[X] \therefore k = 1$

$\phi_X(t) = \frac{2}{5}e^t + \frac{1}{5}e^{2t} + \frac{2}{5}e^{3t}$

$$\begin{aligned}\mathbb{E}[X] &= \lim_{t \to 0} \frac{d\phi_X(t)}{dt} \\ &= \lim_{t \to 0} \left(\frac{2}{5}e^t + \frac{2}{5}e^{2t} + \frac{6}{5}e^{3t} \right) \\ &= \frac{2}{5} + \frac{2}{5} + \frac{6}{5} \\ &= 2\end{aligned}$$

Mean $= 2$

Question 2

(a) Maximum Likelihood Estimator

$X \sim Exp(\lambda) \qquad f(x) = \lambda e^{-\lambda x}$

Let $L(\lambda) = f(x_1, ..., x_n; \lambda)$, $\lambda > 0$ be the joint p.d.f. of $X_1, ..., X_n$.
Let $\hat{\lambda}$ be the Maximum Likelihood Estimate s.t.

$$L(\hat{\lambda}) = f(x_1, ..., x_n; \hat{\lambda}) = \max_{\lambda \in (0,\infty]} f(x_1, ..., x_n; \lambda)$$

and

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

Note: $\forall\lambda\,(L(\lambda) > 0)$ $\therefore$ the log of both sides can be taken.

$$\begin{aligned}\ln L(\lambda) &= \ln\left(\lambda^n e^{-\sum\limits_{i=1}^{n}\lambda x_i}\right)\\ &= n\ln\lambda - \lambda\sum_{i=1}^{n} x_i\end{aligned}$$

Note: $\ln L(\lambda)$ is differentiable and because $ln(x)$ is a strictly increasing function, the value of λ which maximises $L(\lambda)$ also maximises $\ln L(\lambda)$. Find the saddle point of $\ln L(\lambda)$ by setting derivative equal to 0.

$$\begin{aligned}\frac{d\ln L(\lambda)}{d\lambda} &= 0\\ \frac{n}{\lambda} - \sum_{i=1}^{n} x_i &= 0\\ \lambda &= \frac{n}{\sum\limits_{i=1}^{n} x_i}\end{aligned}$$

There is a saddle point at $\lambda = \frac{n}{\sum_{i=1}^{n} x_i}$. To check if this is a local maximum or a local minimum we must check if $\dfrac{d^2\ln L(\lambda)}{d\lambda^2}$ is greater or less than 0 at $\lambda = \frac{n}{\sum_{i=1}^{n} x_i}$.

$$\begin{aligned}\frac{d^2\ln L(\lambda)}{d\lambda^2} &= -\frac{n}{\lambda^2}\\ &= -\frac{n}{\frac{n^2}{\left(\sum_{i=1}^{n} x_i\right)^2}}\\ &= -\frac{\left(\sum\limits_{i=1}^{n} x_i\right)^2}{n}\\ &< 0 \quad \text{(because n > 0)}\end{aligned}$$

The second derivative at $\lambda = \frac{n}{\sum_{i=1}^{n} x_i}$ is less than 0 therefore $L(\lambda)$ is at a local maximum at this point. Our maximum likelihood estimator for λ is $\frac{n}{\sum_{i=1}^{n} x_i}$ which is the inverse of the sample mean $\bar{X}^{-1}$. Thus,

$$\text{Maximum Likelihood Estimate for given set} = \left(\tfrac{10+17+26+14+19+15}{6}\right)^{-1} = \tfrac{6}{101}$$

$$\text{Maximum Likelihood Estimator for } \lambda = \bar{X}^{-1}$$

(b) Average Tuition

$\mu = 18\,273$ $\sigma = 2\,100$ $n = 49$

To find: $P(|\bar{X} - \mu| < 550)$

$\bar{X} \sim N(\mu = 18\,273, \frac{\sigma}{\sqrt{n}} = 300)$

Standardise $\bar{X}$:

$P(\bar{X} > x) = P(Z > z)$ where $Z \sim N(0,1)$ and $z = \dfrac{x - \mu}{\frac{\sigma}{\sqrt{n}}}$

$$\begin{aligned}
P(|\bar{X} - \mu| < 550) &= P(|Z| < 1.8\dot{3}) \\
&= 1 - 2 \cdot P(Z > 1.8\dot{3}) \\
&= 1 - 2 \cdot 0.0333\dot{6} \text{ (interpolated from z-tables)} \\
&= 0.933267 \\
P(|\bar{X} - 18\,273| < 550) &= 93.33\%
\end{aligned}$$

Question 3

Mean Age To Start Walking

Two-tailed test.

$\mu = 12.5$ $n = 18$ $\bar{x} = 12.9$ $\bar{s} = 0.8$ $\alpha = 0.01$

$H_0 : \mu = 12.5$ months

$H_A : \mu \neq 12.5$ months

Reject H_0 if p-value of z-test < 0.01.

$$\begin{aligned} z &= \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} \\ &= \frac{12.9 - 12.5}{\frac{0.8}{\sqrt{18}}} \\ &= 2.1213 \end{aligned}$$

Interpolating from z-tables get

$$\begin{aligned} P(Z > 2.1213) &= 0.01700 + 0.13(0.01659 - 0.01700) \\ &= 0.0169467 \end{aligned}$$

The p-value is equal to $2 \cdot P(Z > 2.1213) = 0.0338934$.

Since the p-value is greater than 0.01, the level of significance, we fail to reject the null hypothesis.

Question 4

(a) Sample Size

$\mu =$ Population mean
$\mu_A =$ Student A sample mean
$\mu_B =$ Student B sample mean

From Central Limit Theorem
$\mu_A \dot{\sim} N(\mu, \frac{\sigma}{\sqrt{50}})$ $\quad$ $\mu_B \dot{\sim} N(\mu, \frac{\sigma}{\sqrt{n}})$

$$\begin{aligned} 3\frac{\sigma}{\sqrt{n}} &= \frac{\sigma}{\sqrt{50}} \\ 3\sqrt{50} &= \sqrt{n} \\ n &= 450 \end{aligned}$$

(b) Traffic Flow

$X \sim Weibull(k = 4, \lambda = 2)$

$F(x) = 1 - e^{-\left(\frac{x}{\lambda}\right)^k}$

$$\begin{aligned}
1 - e^{-\left(\frac{x}{\lambda}\right)^k} &= 0.5 \\
e^{-\left(\frac{x}{\lambda}\right)^k} &= 0.5 \\
-\left(\frac{x}{\lambda}\right)^k &= \ln 0.5 \\
\left(\frac{x}{\lambda}\right)^k &= \ln 2 \\
\left(\frac{x}{2}\right)^4 &= \ln 2 \\
\frac{x^4}{16} &= \ln 2 \\
x^4 &= 16 \ln 2 \\
x &= 1.825 \text{ minutes}
\end{aligned}$$

x = 1 minute 49.5 seconds