24500 HW2

Bin Yu

January 23, 2025

Question 1

Since $X_i \sim \mathcal{N}(\mu, 1)$ and the observations are i.i.d., the maximum likelihood estimator (MLE) for μ is the sample mean:

$$\hat{\mu} = \overline{X} \;=\; \frac{1}{n}\sum_{i=1}^{n} X_i.$$

For $\overline{X}$, $E\overline{X} = \frac{1}{n}\sum_{i=1}^{n} E(Xi) = \mu$, $Var(\overline{X}) = \frac{1}{n^2}\sum_{i=1}^{n} Var(Xi) = \frac{1}{n}$

The distribution of $\overline{X}$ is

$$\overline{X} \;\sim\; N\Big(\mu,\; \frac{1}{n}\Big).$$

$$\sqrt{n}\,\frac{\overline{X}-\mu}{1} \;\sim\; N(0,1).$$

A 95% confidence interval for μ is:

$$\Big[\,\overline{X} \;-\; z_{0.025}\sqrt{\tfrac{1}{n}}\;,\;\; \overline{X} \;+\; z_{0.025}\sqrt{\tfrac{1}{n}}\Big],$$

where $z_{0.025} = 1.96$ is the critical value from the standard normal distribution for a 95% confidence level.

The length L of this interval is

$$L \;=\; 2\times\Big(z_{0.025}\sqrt{\tfrac{1}{n}}\Big) \;=\; 2\;\times 1.96\,\frac{1}{\sqrt{n}} \;=\; \frac{3.92}{\sqrt{n}}.$$

We want this length to be at most 0.5, so

$$\frac{3.92}{\sqrt{n}} \;\le\; 0.5.$$

$$3.92 \;\le\; 0.5\,\sqrt{n}$$

$$\sqrt{n} \;\ge\; \frac{3.92}{0.5}$$

$$n \ge 61.47.$$

Since n must be an integer (the sample size),

$$n \;\ge\; 62.$$

The smallest integer n that satisfies the length requirement is 62.

Question 2

Starting with

$$E\big[(\widehat{\theta}-\theta)^2\big],$$

we can add and subtract $E[\widehat{\theta}]$ inside the parentheses, then expand:

$$(\widehat{\theta}-\theta)^2 \;=\; \big(\widehat{\theta}-E[\widehat{\theta}]+E[\widehat{\theta}]-\theta\big)^2.$$

Taking expectation on both sides and expanding the square:

$$E\big[(\widehat{\theta}-\theta)^2\big] \;=\; E\Big[\big(\widehat{\theta}-E[\widehat{\theta}]\big)^2\Big] \;+\; 2\,E\Big[\big(\widehat{\theta}-E[\widehat{\theta}]\big)\,\big(E[\widehat{\theta}]-\theta\big)\Big] \;+\; E\Big[\big(E[\widehat{\theta}]-\theta\big)^2\Big].$$

Here for each term:

$\big(E[\widehat{\theta}]-\theta\big)$ is a constant (it does not depend on the random variable $\widehat{\theta}$).

So $E\Big[\big(E[\widehat{\theta}]-\theta\big)^2\Big]=\big(E[\widehat{\theta}]-\theta\big)^2$, since $\big(E[\widehat{\theta}]-\theta\big)$ is a constant.

For the term $2E\Big[\big(\widehat{\theta}-E[\widehat{\theta}]\big)\,\big(E[\widehat{\theta}]-\theta\big)\Big]$

$\big(E[\widehat{\theta}]-\theta\big)$ is a constant (not random), so it can be factored out the expectation:

$$2E\Big[\big(\widehat{\theta}-E[\widehat{\theta}]\big)\,\big(E[\widehat{\theta}]-\theta\big)\Big]=2\big(E[\widehat{\theta}]-\theta\big)\,E\big[\widehat{\theta}-E[\widehat{\theta}]\big].$$

By definition,

$$E\big[\widehat{\theta}-E[\widehat{\theta}]\big]=E[\widehat{\theta}]-E[\widehat{\theta}]=0.$$

Hence the entire product is 0:

$$2\big(E[\widehat{\theta}]-\theta\big)\,E\big[\widehat{\theta}-E[\widehat{\theta}]\big]=2\big(E[\widehat{\theta}]-\theta\big)\times 0=0.$$

$E\big[\big(\widehat{\theta}-E[\widehat{\theta}]\big)^2\big]$ is $\mathrm{Var}(\widehat{\theta})$ by definition.

Therefore:

$$E\big[(\widehat{\theta}-\theta)^2\big] \;=\; E\Big[\big(\widehat{\theta}-E[\widehat{\theta}]\big)^2\Big] \;+\; 2\,E\Big[\big(\widehat{\theta}-E[\widehat{\theta}]\big)\,\big(E[\widehat{\theta}]-\theta\big)\Big] \;+\; E\Big[\big(E[\widehat{\theta}]-\theta\big)^2\Big].$$

$$E\big[(\widehat{\theta}-\theta)^2\big] \;=\; \mathrm{Var}(\widehat{\theta}) \;+\; 0+\big(E[\widehat{\theta}]-\theta\big)^2.$$

$$E\big[(\widehat{\theta}-\theta)^2\big] \;=\; \mathrm{Var}(\widehat{\theta}) \;+\; \big(E[\widehat{\theta}]-\theta\big)^2.$$

Question 3

(a)

Definition. A stochastic process $\{N(t)\colon t\ge 0\}$ is called a Poisson process with rate $\lambda>0$ if it satisfies:

1. $N(0) = 0$,
2. For any $s, t \geq 0$, the increment $N(t+s) - N(s)$ is independent of the past (independent increments),
3. For any $s, t \geq 0$,

$$N(t+s) - N(s) \ \sim \ \text{Poisson}\big(\lambda t\big).$$

From the third property above (the number of events in an interval of length t is Poisson(λt)), let X and Y denote the number of the event observed in 10 second and 20 second intervals, then:

$$X_1, \ldots, X_{180} \overset{\text{iid}}{\sim} \text{Poisson}\big(10\lambda\big), \quad Y_1, \ldots, Y_{20} \overset{\text{iid}}{\sim} \text{Poisson}\big(20\lambda\big).$$

This follows because each X_i counts events in a 10-second interval, and each Y_j counts events in a 20-second interval, with an underlying Poisson rate λ per second.

The full likelihood $L(\lambda)$ is:

$$L(\lambda) \ = \ \prod_{i=1}^{180} \Big[e^{-10\lambda} \, \frac{(10\lambda)^{x_i}}{x_i!} \Big] \ \times \ \prod_{j=1}^{20} \Big[e^{-20\lambda} \, \frac{(20\lambda)^{y_j}}{y_j!} \Big].$$

Taking natural logarithms:

$$\ell(\lambda) \ = \ \ln L(\lambda) \ = \ \sum_{i=1}^{180} \Big[-10\lambda + X_i \ln(10\lambda) - \ln(X_i!) \Big] \ + \ \sum_{j=1}^{20} \Big[-20\lambda + Y_j \ln(20\lambda) - \ln(Y_j!) \Big].$$

Let

$$S_X \ = \ \sum_{i=1}^{180} X_i, \qquad S_Y \ = \ \sum_{j=1}^{20} Y_j.$$

Then we can rewrite $\ell(\lambda)$ as:

$$\ell(\lambda) \ = \ -10\lambda \times 180 \ - \ 20\lambda \times 20 \ + \ S_X \ln(10\lambda) \ + \ S_Y \ln(20\lambda) \ + \ \big(\text{terms independent of } \lambda\big).$$

So

$$\ell(\lambda) \ = \ -2200\,\lambda \ + \ S_X \ \ln(10\lambda) \ + \ S_Y \ \ln(20\lambda) \ + \ \big(\text{terms independent of } \lambda\big).$$

$$\frac{d}{d\lambda}\,\ell(\lambda) = -2200 \ + \ S_X \cdot \frac{d}{d\lambda}[\ln(10\lambda)] \ + \ S_Y \cdot \frac{d}{d\lambda}[\ln(20\lambda)].$$

$$\frac{d}{d\lambda}\big[\ln(10\lambda)\big] \ = \ \frac{1}{10\lambda} \cdot 10 \ = \ \frac{1}{\lambda}, \quad \frac{d}{d\lambda}\big[\ln(20\lambda)\big] \ = \ \frac{1}{20\lambda} \cdot 20 \ = \ \frac{1}{\lambda}.$$

Thus

$$\frac{d}{d\lambda}\,\ell(\lambda) = -2200 \ + \ \frac{S_X}{\lambda} \ + \ \frac{S_Y}{\lambda} \ = \ -2200 \ + \ \frac{S_X + S_Y}{\lambda}.$$

Setting this derivative to zero:

$$-2200 \ + \ \frac{S_X + S_Y}{\lambda} \ = \ 0$$

$$\widehat{\lambda} \ = \ \frac{S_X + S_Y}{2200}.$$

Now substitute the data:

$$S_X = \sum_{i=1}^{180} X_i = 0 \times 23 \ + \ 1 \times 77 \ + \ 2 \times 34 \ + \ 3 \times 26 \ + \ 4 \times 13 \ + \ 5 \times 7 = 310.$$

$$S_Y = \sum_{i=1}^{20} Y_i = 0 \times 2 \;+\; 1 \times 4 \;+\; 2 \times 9 \;+\; 3 \times 5 \;=\; 37.$$

$$S_X = 310, \quad S_Y = 37, \quad \widehat{\lambda} = \frac{310 + 37}{2200} \approx 0.1577$$

(b)

Suppose we have an i.i.d. sample $X_1, X_2, \ldots, X_n \sim \text{Poisson}(\lambda)$. The pmf of each single observation X_i is

$$P(X_i = x_i) \;=\; \frac{\lambda^{\,x_i}\, e^{-\lambda}}{x_i!}.$$

the likelihood is

$$L(\lambda) \;=\; \prod_{i=1}^{n} \Big[e^{-\lambda}\, \frac{\lambda^{x_i}}{x_i!} \Big] \;=\; e^{-n\lambda}\, \frac{\lambda^{\sum_{i=1}^n x_i}}{\prod_{i=1}^n x_i!}.$$

Taking log:

$$\ell(\lambda) = -\,n\,\lambda \;+\; \Big(\sum_{i=1}^{n} x_i\Big)\; \ln(\lambda) \;-\; \sum_{i=1}^{n} \ln\big(x_i!\big).$$

Set the derivative to 0:

$$\frac{d}{d\lambda}\,\ell(\lambda) = -\,n \;+\; \frac{1}{\lambda} \sum_{i=1}^{n} x_i \;\;=\; 0$$

$$\widehat{\lambda} = \frac{1}{n} \sum_{i=1}^{n} x_i = \overline{X}.$$

Note that the sum of all 200 counts has a Poisson distribution:

$$\sum_{i=1}^{180} X_i \;+\; \sum_{i=1}^{20} Y_i \;\sim\; \text{Poisson}\big(2200\,\lambda\big) \;\equiv\; \sum_{i=1}^{2200} Z_i,$$

where each $Z_i \overset{\text{iid}}{\sim} \text{Poisson}(\lambda)$.

Since Z_i are i.i.d. Poisson(λ), each has mean $E[Z_i] = \lambda$ and variance $\text{Var}(Z_i) = \lambda$. Then

$$Z = \sum_{i=1}^{2200} Z_i \quad \text{has} \quad E[Z] = 2200\,\lambda, \quad \text{Var}(Z) = 2200\,\lambda.$$

By CLT, Z can be approximated by a normal distribution:

$$Z \;\approx\; N\big(2200\,\lambda,\; 2200\,\lambda\big).$$

$$\frac{Z - 2200\,\lambda}{\sqrt{2200\,\lambda}} \;\xrightarrow{\text{d}}\; N(0,1).$$

the MLE for Z is

$$\widehat{\lambda} \;=\; \frac{1}{2200} \sum_{i=1}^{2200} Zi = \frac{Z}{2200}.$$

Since $Z \approx N(2200\,\lambda,\, 2200\,\lambda)$, dividing by 2200 gives

$$\widehat{\lambda} \;=\; \frac{Z}{2200} \;\approx\; N\Big(\lambda,\; \frac{\lambda}{2200}\Big).$$

Therefore,

$$\widehat{\lambda} \;\approx\; N\Big(\lambda,\; \frac{\lambda}{2200}\Big).$$

Question 4

(a)

Let Z be a standard normal random variable, i.e. $Z \sim N(0,1)$.

$$E[Z^2] \;=\; 1, \qquad E[Z^4] \;=\; 3.$$

(b)

Let $Z_1, Z_2, \ldots, Z_n$ be i.i.d. $N(0,1)$ random variables, and define

$$Y \;=\; \sum_{i=1}^{n} Z_i^2.$$

For mean of Y, since each Z_i^2 has expectation

$$E[Z_i^2] \;=\; 1$$

because $Z_i \sim N(0,1)$. Therefore,

$$E[Y] \;=\; E\Big[\sum_{i=1}^{n} Z_i^2\Big] \;=\; \sum_{i=1}^{n} E[Z_i^2] \;=\; n \cdot 1 \;=\; n.$$

For variance of Y, since $Z_1^2, Z_2^2, \ldots, Z_n^2$ are independent,

$$\mathrm{Var}\Big(\sum_{i=1}^{n} Z_i^2\Big) \;=\; \sum_{i=1}^{n} \mathrm{Var}(Z_i^2).$$

for a standard normal variable Z_i,

$$\mathrm{Var}(Z_i^2) \;=\; E[Z_i^4] \;-\; \big(E[Z_i^2]\big)^2.$$

We know $E[Z_i^2] = 1$ and $E[Z_i^4] = 3$,

$$\mathrm{Var}(Z_i^2) \;=\; 3 - 1^2 \;=\; 2.$$

Hence,

$$\mathrm{Var}(Y) \;=\; \sum_{i=1}^{n} \mathrm{Var}(Z_i^2) \;=\; \sum_{i=1}^{n} 2 \;=\; 2n.$$

Therefore,

$$E[Y] \;=\; n \quad \text{and} \quad \mathrm{Var}(Y) \;=\; 2n.$$