[I225] Statistical Signal Processing(E) Office Hour 4

1. There are three machines, A, B, and C, that manufacture a certain product. Machines A, B, and C produce 20%, 30%, and 50% of the total products, respectively. It is known from experience that 5%, 4%, and 2% of the products from machines A, B, and C, respectively, are defective.

a) What is the probability that a randomly selected product from the total production is defective?

b) Given that a product is found to be defective, what is the probability that it was produced by machine A, B, or C?

Answer:

Let the events that a randomly selected product was produced by machines A, B, and C be denoted by *A*, *B*, and *C*, respectively.
Let the event that the selected product is defective be denoted by E.

$$\mathrm{P(A) = 0.2,\ P(B) = 0.3,\ P(C) = 0.5}$$

$$\mathrm{P(E|A) = 0.05,\ P(E|B) = 0.04,\ P(E|C) = 0.02}$$

Since A, B, and C are mutually exclusive and exhaustive events:

$$\mathrm{P(E) = P(A)P(E|A) + P(B)P(E|B) + P(C)P(E|C) = 0.032}$$

Also, by Bayes' theorem:

$$\mathrm{P(A|E)} = \frac{\mathrm{P(A)P(E|A)}}{\mathrm{P(E)}} = \frac{5}{16},\ \mathrm{P(B|E)} = \frac{\mathrm{P(B)P(E|B)}}{\mathrm{P(E)}} = \frac{3}{8}$$

$$\mathrm{P(C|E)} = \frac{\mathrm{P(C)P(E|C)}}{\mathrm{P(E)}} = \frac{5}{16}$$

So the final answers are:

The probability that a randomly selected product is defective: 0.032

The probabilities that the defective product came from each machine:

$$\mathrm{P(A|E)} = \frac{5}{16},\ \mathrm{P(B|E)} = \frac{3}{8},\ \mathrm{P(C|E)} = \frac{5}{16}$$

2. In a lottery, 6 numbers are drawn from a pool of 33 red balls numbered from 1 to 33. If all selected numbers match the drawn numbers, the player wins the second prize.

Suppose the probability of winning the second prize with a single ticket is

$$p = \frac{1}{\binom{33}{6}} \approx 9.0288 \times 10^{-7}$$

Assume the following:

- The lottery is held 3 times per week,
- The player buys 10 tickets per draw,
- The player keeps buying tickets for 5 years (52 weeks per year).

Question: What is the probability that the player wins the second prize at least once during the 5 years?

Answer:

For simplicity, assume there are 52 weeks in a year, and the lottery is drawn 3 times a week, with 10 tickets purchased each time. Then, over 5 years, the total number of tickets purchased is:

$$52 \times 3 \times 10 \times 5 = 7800$$

Let X be the number of times the person wins the second prize. Then:

$$X \sim B(7800, p)$$

According to Poisson approximation of the binomial distribution, since:

$$\lambda = np = 7800 \times 9.0288 \times 10^{-7} = 0.007042$$

The probability of winning at least once is:

$$P(X \geq 1) = 1 - P(X = 0) = 1 - (1 - p)^{7800} \approx 1 - \exp(-0.007042) = 1 - 0.993$$
$$= 0.007$$

So, the probability of winning the second prize at least once in 5 years is approximately 0.007, which is considered a rare event under strict statistical standards.

3. Let X be a random variable uniformly distributed on the interval $(0,1)$, and let Y be a random variable uniformly distributed on $(0, X)$. Find the probability density function (PDF) of Y, and compute the conditional cumulative distribution function $F_{X|Y}(0.5|0.25)$.

Answer:

From the problem, we know: $X \sim U(0|1), Y|X = x \sim U(0, x)$

Based on Conditional probability density function (If the probability density function of Y at the point y, denoted $f_2(y)$, satisfies $f_2(y) > 0$, then the conditional probability density function of X given $Y = y$ is defined as: $f_{X|Y}(x|y) = \frac{f(x,y)}{f_2(y)}$).

The joint density function of (X, Y) is:

$$f(x, y) = f_1(x) f_{Y|X}(y|x) = I_{(0,1)}(x) \cdot \frac{1}{x} \cdot I_{(0,x)}(y)$$

For $0 < y < 1$, the marginal density of Y is:

$$f_2(y) = \int_{-\infty}^{\infty} f(x, y)\, dx = \int_0^1 \frac{1}{x} \cdot I_{(y<x)}\, dx = \int_y^1 \frac{1}{x} dx = -\ln(y) \cdot I_{(0,1)}(y)$$

From the definition of the conditional density, we have:

$$f_{X|Y}(x|0.25) = \frac{f(x, 0.25)}{f_2(0.25)} = \frac{1/x}{\ln 4} \cdot I_{(0.25,1)}(x) = \frac{1}{x \ln 4} \cdot I_{(0.25,1)}(x)$$

Therefore, the conditional CDF is:

$$f_{X|Y}(0.5|0.25) = \int_{-\infty}^{0.5} f_{X|Y}(x|0.25)\, dx = 0.5$$

4. A call center records the number of calls received during 5 different one-hour intervals: Data: [3, 2, 4, 3, 2]

Assume the number of calls received in one hour follows a Poisson distribution with rate λ (calls per hour), and each hour is independent.

Answer the following:

a) Write the likelihood function for this data given the parameter λ

b) Write the log-likelihood function.

c) Find the maximum likelihood estimate $\hat{\lambda}$ for the rate of calls per hour.

Answer:

a) Likelihood Function:

The Poisson distribution's probability mass function is:

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

For independent samples $x_1, x_2, x_3, \ldots, x_n$, the likelihood is:

$$L(\lambda) = \prod_{i=1}^{n} \frac{\lambda^{x_i} e^{-\lambda}}{x_i!} = \frac{\lambda^{\sum x_i} e^{-n\lambda}}{\prod x_i!}$$

In our data [3, 2, 4, 3, 2], we have:

$$\sum x_i = 14,$$

$$n = 5$$

So

$$L(\lambda) = \frac{\lambda^{14} e^{-5\lambda}}{3! \cdot 2! \cdot 4! \cdot 3! \cdot 2!}$$

We write as

$$L(\lambda) = \frac{\lambda^{14} e^{-5\lambda}}{C}, \text{where } C = \prod x_i!$$

b) Log-Likelihood Function

Take the logarithm:

$$\ell(\lambda) = log\mathrm{L}(\lambda) = 14log\lambda - 5\lambda + \mathrm{const}$$

c) Maximum Likelihood Estimate

Differentiate and set to zero:

$$\frac{d\ell}{d\lambda} = \frac{14}{\lambda} - 5 = 0 \rightarrow \frac{14}{\lambda} = 5$$

$$\lambda = 2.8$$

5. A researcher is studying the lifetimes of batteries used in remote sensors. Suppose the lifetimes are modeled as a vector

$$L = [L[0], L[1], \ldots, L(N-1)]^T,$$

where each observation is an independent and identically distributed sample from an exponential distribution with unknown parameter λ (the rate parameter):

$$L[n] \sim \mathrm{Exponential}(\lambda), \mathrm{n} = 0, \ldots, N-1$$

Find the maximum likelihood estimator (MLE) $\hat{\lambda}$ for λ.

Answer:

The probability density function (pdf) of the exponential distribution is:

$$f(;\lambda) = \lambda e^{-\lambda x}, x \geq 0$$

Step 1: Write the likelihood function

$$L(\lambda) = \prod_{n=0}^{N-1} \lambda e^{-\lambda L[n]} = \lambda^N e^{-\lambda \sum_{n=0}^{N-1} L[n]}$$

Step 2: Log-likelihood

$$logL(\lambda) = Nlog\lambda - \lambda \sum_{n=0}^{N-1} L[n]$$

Step 3: Take derivative and set to zero

$$\frac{d}{d\lambda} logL(\lambda) = \frac{N}{\lambda} - \sum_{n=0}^{N-1} L[n] = 0$$

$$\rightarrow \hat{\lambda} = \frac{N}{\sum_{n=0}^{N-1} L[n]}$$

That is, the MLE of λ is the reciprocal of the sample mean.