EE 3070 Statistics

Midterm Exam Solution

April 11, 2023
10:10 ∼ 12:00

Note: There are **8** problems with total 100 points within **2** pages, please write your answer with detail in the answer sheet.
No credit without detail, except for question 1. No calculator. Closed books.

1. (12%) We observed $x_1, x_2, \cdots, x_{100}$ independent samples from a Gaussian random variable $N(\mu, \sigma^2)$, where μ and σ^2 are unknown. Denote $\bar{x} = \frac{\sum\limits_{i=1}^{100} x_i}{100}$. Which of the following statements is **TRUE**?

    (a) The outcome from the estimator (for μ) $\widehat{x}_A = \frac{\sum\limits_{i=1}^{100} x_i}{100}$ will definitely be more accurate than $\widehat{x}_B = \frac{\sum\limits_{i=1}^{50} x_i}{50}$.

    (b) The estimator (for μ) $\widehat{x}_C = \frac{x_1 + x_2}{2}$ is a valid and unbiased estimator.

    (c) The estimator (for σ^2) $\widehat{\sigma_A^2} = \frac{\sum\limits_{i=1}^{100} (x_i - \bar{x})}{100}$ is the maximum likelihood and unbiased estimator.

    (d) Assuming σ^2 is known, the likelihood function $f(x_i \, ; \mu)$ has the property $\int f(x_i \, ; \mu) \, d\mu = 1$.

    (e) For any point in the parameter space (μ, σ^2), the likelihood function $f(x_i \, ; \, \mu, \sigma^2) \leq 1$.

    Solution: (b), (d), (e)

    Without any prior knowledge, option (d) and (e) cannot be true because the likelihood function is not a probability function. There is no guarantee that the integration of $\int f(x_i \, ; \mu) \, d\mu$ will integrate to 1 over all possible values of μ.

    Since in our case, x_i is a sample from a Gaussian random variable, and we know the integral of the PDF over the entire range of the outcomes is equal to 1.

    $$\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(x_i - \mu)^2}{2\sigma^2}} \, dx_i = 1 \iff \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \cdot e^{-\frac{(\mu - x_i)^2}{2\sigma^2}} \, d\mu = 1$$

    Therefore, in this case, $\int f(x_i \, ; \mu) \, d\mu$ will also equal to 1. But please be reminded that this is not always the case. The likelihood function has no such property that integrates to 1 with respect to any parameter.

    ∎

2. (12%) Assumed that discrete random variable X has a moment generating function given by

$$M_X(t) = \frac{1}{6} \cdot e^{-2t} + \frac{1}{3} \cdot e^{-t} + \frac{1}{8} \cdot e^{t} + \frac{3}{8} \cdot e^{2t}$$

Please find $P(|X| \leq 1)$, $E(X)$, $Var(X)$.

Solution:

$$M_X(t) = E(e^{tx}) = \sum_{n=0}^{\infty} e^{tx} \cdot P_X(x)$$

$$P(X = 1) = \frac{1}{8}, \ P(X = 2) = \frac{3}{8}, \ P(X = -1) = \frac{1}{3}, \ P(X = -2) = \frac{1}{6}$$

$$P(|X| \leq 1) = P(X = 1) + P(X = -1) = \frac{1}{8} + \frac{1}{3} = \frac{11}{24}$$

$$E(X) = \frac{5}{24}, \ E(X^2) = \frac{21}{8}, \ Var(X) = \frac{21}{8} - (\frac{5}{24})^2 = \frac{1487}{576}$$

■

3. (12%) A random variable X has the Poisson distribution given by

$$p(x; \mu) = \frac{e^{-\mu} \mu^x}{x!}, \quad \text{for } x = 0, 1, 2, \cdots$$

(a) Please derive the moment generating function of X.
(b) Please use the MGF in part (a) to find the mean and the variance of X.
Hint: Let $K(t) = \ln M_X(t) = ?$

Solution:
(a) To derive $M_X(t)$, we can use the definition of the moment generating function:

$$\sum_{x=0}^{\infty} \frac{e^{-\mu} \mu^x}{x!} e^{tx} = \sum_{x=0}^{\infty} \frac{(e^t \mu)^x}{x!} e^{-\mu} = e^{-\mu} e^{e^t \mu} = e^{\mu(e^t - 1)}$$

(b) To find the mean and the variance of X, we can first compute $K(t) = \ln M_X(t)$.

$$K(t) = \ln(M_X(t)) = \mu(e^t - 1)$$

Therefore, the mean and the variance of X are given by:

$$\mu = k'(t)\big|_{t=0} = \mu$$

$$\sigma^2 = k''(t)\big|_{t=0} = \mu$$

■

2

4. (13%) Consider the mixture distribution, $\frac{7}{10} \cdot N(0,4) + \frac{3}{10} \cdot N(0,16)$. Find the kurtosis.

Solution:

In our case, we have a finite and countable mixtures, which means given a finite set of PDFs $f_1(x), ..., f_n(x)$ and weights $w_1, ..., w_n$ such that $w_i \geq 0$ and $\sum w_i = 1$, the mixture distribution can be represented by writing the density function, f, as a sum (which is a convex combination):

$$f(x) = \sum_{i=1}^{n} w_i \cdot f_i(x)$$

To find the kurtosis, we first need to find the variance of this mixture distribution. In this case of a mixture of one-dimensional distributions with weights w_i, mean μ_i and variances σ_i^2, the total mean and variance will be:

$$E(X) = \mu = \sum_{i=1}^{n} w_i \cdot \mu_i$$

$$
\begin{aligned}
E((X-\mu)^2) &= \sigma^2 \\
&= E(X^2) - \mu^2 && \text{(standard variance reformulation)} \\
&= \left(\sum_{i=1}^{n} w_i \cdot E(X_i^2) \right) - \mu^2 \\
&= \left(\sum_{i=1}^{n} w_i \cdot (\sigma_i^2 + \mu_i^2) \right) - \mu^2 \quad (\because \sigma_i^2 = E(X_i^2) - \mu_i^2 \quad \therefore E(X_i^2) = \sigma_i^2 + \mu_i^2)
\end{aligned}
$$

Applying the above equations to our case gives us the following results for the mean and variance.

Note that although X represents a random variable from a mixture distribution of two Gaussian distributions, we cannot assume that X is also a Gaussian variable, even though we can compute its mean and variance using a simple weighted sum. Later on, we will find out that X is truly not Gaussian.

$$f_1(x) \sim N(0,4), \quad f_2(x) \sim N(0,16), \quad f(x) = \frac{7}{10} \cdot f_1(x) + \frac{3}{10} \cdot f_2(x)$$

$$E(X) = \frac{7}{10} \cdot 0 + \frac{3}{10} \cdot 0 = 0, \quad \text{Var}(X) = \frac{7}{10} \cdot 4 + \frac{3}{10} \cdot 16 = \frac{76}{10} = 7.6 = \sigma_x^2$$

We know that in order to compute kurtosis, we first need to find $E(X^4)$. Let's start with deriving the formula for $E(X^4)$ in the general case, regardless of the distribution of X. Since we already know that if X is Gaussian, then $E(X^4)$ is equal to $3\sigma_x^4$.

We can find the formula for $E(X^4)$ by using the definition:

$$
\begin{aligned}
E(X^4) &= \int_{-\infty}^{\infty} x^4 \cdot f_X(x) \, dx = \int_{-\infty}^{\infty} x^4 \cdot \left(\sum_{i=1}^{n} w_i \cdot f_i(x) \right) dx = \sum_{i=1}^{n} w_i \cdot \left(\int_{-\infty}^{\infty} x^4 \cdot f_i(x) \, dx \right) \\
&= \sum_{i=1}^{n} w_i \cdot E(X_i^4)
\end{aligned}
$$

$$\Rightarrow E(X^4) = \frac{7}{10} \cdot E(X_1^4) + \frac{3}{10} \cdot E(X_2^4) = 264 \neq 173.28 = 3 \times 7.6^2 = 3\sigma_x^4 \quad \therefore X \text{ is not Gaussian.}$$

Then, we continued with the definition of kurtosis:

$$k = \frac{E(X^4)}{\text{Var}(X)^2} = \frac{\frac{7}{10} \cdot E(X_1^4) + \frac{3}{10} \cdot E(X_2^4)}{(\frac{76}{10})^2} = \frac{0.7 \times 3 \times 2^4 + 0.3 \times 3 \times 4^4}{76^2/10^2} = \frac{264 \cdot 100}{5776}$$

$$= \frac{1}{10} \cdot \frac{7 \cdot (3 \cdot \sigma_1^4) + 3 \cdot (3 \cdot \sigma_2^4)}{76^2} \cdot 10^2$$

$$= \frac{21 \cdot 2^4 + 9 \cdot 4^4}{76^2} \cdot 10 = \frac{21 \cdot 16 + 9 \cdot 256}{5776} \cdot 10$$

$$= \frac{26400}{5776} = \frac{1650}{361} \approx 4.57063712$$

$\blacksquare$

5. (13%) Suppose the number of customers X that enter a store between the hours 8:00 a.m. and 9:00 a.m. follows a Poisson distribution with parameter θ. Suppose a random sample of the number of customers that enter the store between 8:00 a.m. and 9:00 a.m. for 8 days results in the values: 4, 8, 12, 13, 9, 6, 5, 12.
(a) Determine the maximum likelihood estimate of θ. Show that it is an unbiased estimator.
(b) Based on these data, obtain the realization of your estimator in part (a). Explain the meaning of this estimate in terms of the number of customers.

Solution:
(a) To compute $\widehat{\theta}$, we can use the definition of the ML estimator:

$$X \sim P(\theta) \quad P_X(x) = \frac{e^{-\theta} \cdot \theta^x}{x!}, \quad x = 0, 1, 2, \cdots$$

$$L(\theta) = \prod_{i=1}^{n} P(x_i > \theta) = \prod_{i=1}^{n} \frac{e^{-\theta} \cdot \theta^{x_i}}{x_i!} = \frac{e^{-n\theta} \cdot \theta^{x_1+x_2+\cdots+x_n}}{x_1!x_2!\cdots x_n!}$$

$$\Rightarrow \ln(L(\theta)) = \ln(e^{-n\theta} \cdot \theta^{x_1+x_2+\cdots+x_n}) - \ln(x_1!x_2!\cdots x_n!)$$

$$= \ln(e^{-n\theta}) + \ln(\theta^{x_1+x_2+\cdots+x_n}) + \ln(x_1!x_2!\cdots x_n!)$$

$$= -n \cdot \theta + (\sum_{i=1}^{n} x_i) \cdot \ln(\theta) - \ln(\prod_{i=1}^{n} x_i!)$$

$$\Rightarrow \frac{\partial \ln(L(\theta))}{\partial \theta} = -n + \frac{1}{\theta} \sum_{i=1}^{n} x_i = 0$$

$$\Rightarrow \frac{1}{\theta} \sum_{i=1}^{n} x_i = n$$

$$\Rightarrow \theta = \frac{1}{n} \sum_{i=1}^{n} x_i = \bar{x}$$

$$\Rightarrow \text{ML estimator} \quad \widehat{\theta} = \bar{x}$$

To show this estimator to be unbiased, we need to find its expected value of $\bar{x}$ and show that it is equal to θ. The expected value of $\bar{x}$ is:

$$E(\bar{x}) = E(\frac{1}{n} \sum_{i=1}^{n} x_i) = \frac{1}{n} \cdot E(\sum_{i=1}^{n} x_i) = \frac{1}{n} \cdot \sum_{i=1}^{n} \theta = \frac{1}{n} \cdot n \cdot \theta = \theta$$

4

Therefore, we have shown that the ML estimator $\widehat{\theta} = \bar{x}$ is unbiased.

(b)

$$\hat{\theta} = \frac{1}{8}(4 + 8 + 12 + 13 + 9 + 6 + 5 + 12) = \frac{69}{8} = 8.625$$

∎

6. (13%) Let X, Y, and Z be three independent Poisson random variables with parameters $\lambda_1, \lambda_2, \lambda_3$, respectively. For $y = 0, 1, 2, \cdots, t$, calculate $E(Y|X + Y + Z = t)$.

Solution:
To compute $E(Y|X + Y + Z = t)$, we can use the definition of conditional probability:

$$
\begin{aligned}
P(Y|X + Y + Z = t) &= \frac{P(Y, \ X + Y + Z = t)}{P(X + Y + Z = t)} = \frac{P(Y)P(X + Z = t - Y)}{P(X + Y + Z = t)} \\
&= P(Y) \cdot P(X + Z = t - Y) \cdot \frac{1}{P(X + Y + Z = t)} \\
&= \frac{e^{-\lambda_2} \cdot \lambda_2^y}{y!} \cdot \frac{e^{-(\lambda_1 + \lambda_3)} \cdot (\lambda_1 + \lambda_3)^{t-y}}{(t - y)!} \cdot \frac{t!}{e^{-(\lambda_1 + \lambda_2 + \lambda_3)} \cdot (\lambda_1 + \lambda_2 + \lambda_3)^t} \\
&= \frac{(\lambda_1 + \lambda_3)^{t-y} \cdot \lambda_2^y \cdot t!}{(\lambda_1 + \lambda_2 + \lambda_3)^t \cdot y!(t - y)!} \\
&= \frac{t!}{y!(t - y)!} \cdot \frac{(\lambda_1 + \lambda_3)^{t-y} \cdot \lambda_2^y}{(\lambda_1 + \lambda_2 + \lambda_3)^y} \\
&= C_y^t \cdot \left(\frac{\lambda_2}{\lambda_1 + \lambda_2 + \lambda_3}\right)^y \cdot \left(\frac{\lambda_1 + \lambda_3}{\lambda_1 + \lambda_2 + \lambda_3}\right)^{t-y} \\
&\sim Bernoulli(t, \frac{\lambda_2}{\lambda_1 + \lambda_2 + \lambda_3})
\end{aligned}
$$

Therefore,

$$E(Y|X + Y + Z = t) = t \cdot \frac{\lambda_2}{\lambda_1 + \lambda_2 + \lambda_3}$$

∎

7. (12%) Let $X_1, X_2, ..., X_N$ be a set of independent random variables, where each X_i is a normal random variable with mean equal to μ and variance equal to σ^2. Please derive the moment generating function of Y, where $Y = X_1 + X_2 + \cdots + X_N$ and N is Poisson random variable with mean λ.

Solution:
From the problem description, we know that:

$$X_i \sim N(\mu, \sigma^2) \ ; \ N \sim Poisson(\lambda) \ ; \ Y = X_1 + \cdots + X_N$$

To derive the mgf of Y, we start with the definition:

$$M_Y(t) = E(e^{ty}) = E\left(E(e^{ty}|N)\right)$$

$$= E\left(E(e^{t(X_1+X_2+\cdots+X_N)|N})\right) = E\left((e^{\mu t+\frac{1}{2}\sigma^2 t^2})^N\right)$$

$$= \sum_{n=0}^{\infty}(e^{\mu t+\frac{1}{2}\sigma^2 t^2})^N = n \cdot P_N(n)$$

$$= \sum_{n=0}^{\infty} e^{n\mu t+\frac{1}{2}n\sigma^2 t^2} \cdot \left(\frac{e^{-\lambda}\lambda^n}{n!}\right)$$

$$= e^{-\lambda} \cdot \left(\sum_{n=0}^{\infty}\frac{1}{n!} \cdot (\lambda \cdot e^{\mu t+\frac{1}{2}\sigma^2 t^2})^n\right)$$

$$= e^{-\lambda} \cdot \left(e^{\lambda \cdot (e^{\mu t+\frac{1}{2}\sigma^2 t^2})}\right) = e^{\lambda(e^{\mu t+\frac{1}{2}\sigma^2 t^2}-1)}$$

Therefore, the mgf of Y is given by

$$M_Y(t) = \exp\{\lambda \cdot (\exp\{\mu t + \frac{1}{2}\sigma^2 t^2\} - 1)\}$$

∎

8. (13%) Consider a signal

$$x[n] = A \cdot \cos(2\pi f_0 n + \phi) + w[n], \ n = 0, 1, \cdots, N-1.$$

The A and f_0 are assumed known. The n is the sampling time index. The $w[n]$ is distributed as $N(0, \sigma^2)$, where σ^2 is known. The $x[n]$ is the observed data. The ϕ is the parameter to be found.
(a) Derive the likelihood function of ϕ.
(b) Derive the procedure to find the Maximum Likelihood Estimation of ϕ.

Solution:
(a) The likelihood function of ϕ is given by:

$$P(x; \phi) = \frac{1}{(2\pi\sigma^2)^{\frac{N}{2}}} \cdot e^{-\frac{1}{2\sigma^2}\sum_{n=0}^{N-1}(x[n]-A\cos(2\pi f_0 n+\phi))}.$$

(b) Take the first order derivative of the log likelihood function to obtain:

$$\frac{\partial \ln\{p(x; \phi)\}}{\partial \phi} = \cdots = -\frac{A}{\sigma^2}\sum_{n=0}^{N-1}\left(x[n]\sin(2\pi f_0 n + \phi) - \frac{A}{2}\sin(4\pi f_0 n + 2\phi)\right)$$

Set $\frac{\partial \ln\{p(x; \phi)\}}{\partial \phi} = 0$ to solve for ϕ (no closed form)

∎