Probability & Statistics for EECS: Homework #13

Due on May 14, 2023 at 23:59

Name: **Wang Yunfei**
Student ID: 2021533135

Problem 1

Solution

(a) First for any X_j, we have,

$$\begin{aligned} P(X_j >= 1) &= P(X_j > 1) \\ &= 1 - P(X_j <= 1) \\ &= 1 - (1 - \frac{1}{e}) \\ &= \frac{1}{e} \end{aligned}$$

And then from the question, N is the r.v. which represents the fisrt X_j exceeds 1.
Therefore we can seem N as the Fisrt success distribution, and the parameter $p = \frac{1}{e}$.
Then from what we have learned, we know N-1 actually is distributed as Geometric distribution, that is $N - 1 \sim Geom(\frac{1}{e})$.
At the meanwhile, $E[N] = p = e$.

(b) First in this question, we should think of the relationship between Poisson distribution and Exponential distribution. From what we have learned, the r.v. M-1 can be seemed as during the time ranging from 0 to 10, how many arrivals occurs in this period.
And because $X_1, X_2, \cdots \sim Expo(1)$ and $t = 10$. Then $\lambda t = 10$
Then this question actually is the number of arrivals that occurs in an interval of length 10, that is $M - 1 \sim Pois(10)$, because M actually has exceeded or been equal to 10.
Then $E[M - 1] = E[M] - 1 = 10$.
Therefore $E[M] = 11$.

(c) First, for $\frac{X_j}{n}$, let $Y = \frac{X_j}{n}$. Then $P(Y <= x) = P(X_j <= nx) = 1 - e^{-nx}$. Therefore we can get $Y = \frac{X_j}{n} \sim Expo(n)$.
Second, from the theorem we learned about Gamma distribution, if $X_1, \cdots X_n \sim Expo(\lambda)$ and they are i.i.d.s, then we can have $X_1 + \cdots + X_n \sim Gamma(n, \lambda)$.
Then we can have $\overline{X_n} \sim Gamma(n, n)$.
And from the properties of Gamma distribution, we have $E[\overline{X_n}] = 1, Var[\overline{X_n}] = \frac{1}{n}$. And from the properties of Exponential distribution, we have $E[X_j] = \frac{1}{\lambda} = 1 = \mu, Var[X_j] = \frac{1}{\lambda^2} = 1 = \sigma^2$.
Then we have $E[\overline{X_n}] = \mu, Var[\overline{X_n}] = \frac{\sigma^2}{n}$.
Then we can use CLT to get the right answer, that is when $n \to \infty$, $\overline{X_n} \sim N(1, \frac{1}{n})$.

Problem 2

(a) Denote $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$.
Let $S_n = \sum_{i=1}^{n} X_i$ and when $s > 0, t > 0$, we can get the following formula from chernoff's ineuqality,

$$\begin{aligned} P(S_n - E[S_n] \geq t) &\leq e^{-st} E[e^{s(S_n - E[S_n])}] \\ &= e^{-st} \prod_{i=1}^{n} E[e^{s(X_i - E[X_i])}] \end{aligned}$$

$$from\ the\ Hoeffding\ Lemma\ ,\ we\ can\ have$$

$$\begin{aligned} P(S_n - E[S_n] \geq t) &\leq e^{-st} \prod_{i=1}^{n} e^{\frac{s^2(b_i - a_i)^2}{8}} \\ &= e^{-st + \frac{1}{8}s^2 \sum_{i=1}^{n}(b_i - a_i)^2} \end{aligned}$$

Because we need to find the min value of $e^{-st+\frac{1}{8}s^2\sum_{i=1}^{n}(b_i-a_i)^2}$, then let $f(s) = -st + \frac{1}{8}s^2\sum_{i=1}^{n}(b_i - a_i)^2$, and we need to find the the min of $f(s)$.
Obviously $f^2(x) >= 0$, then we just need to find $f'(s) = 0$.
Finally we can get $s = \frac{4t}{\sum_{i=1}^{n}(b_i-a_i)^2}$, and then substitute into equation, we can get

$$P(S_n - E[S_n] \geq t) \leq e^{\frac{-2t^2}{\sum_{i=1}^{n}(b_i-a_i)^2}}$$

And because $\overline{X} = \frac{S_n}{n}$, then we can get,

$$P(\overline{X} - E[\overline{X}] \geq t) \leq e^{\frac{-2n^2t^2}{\sum_{i=1}^{n}(b_i-a_i)^2}}$$

Similarly we can get $P(E[\overline{X}] - \overline{X} \geq t) \leq e^{\frac{-2n^2t^2}{\sum_{i=1}^{n}(b_i-a_i)^2}}$.
Then we can get $P(\left|\overline{X} - E[\overline{X}]\right| \geq t) \leq e^{\frac{-2n^2t^2}{\sum_{i=1}^{n}(b_i-a_i)^2}}$.
And because $X_1, X_2, \cdots, X_n$ are independent with $E[X_I] = \mu$, $a \leq X_i \leq b$ for $i = 1, \cdots, n$.
Therefore we can get the final answer,

$$\begin{aligned} P(\overline{X} - E[\overline{X}] \geq t) &\leq e^{\frac{-2n^2t^2}{n(b_i-a_i)^2}} \\ &\leq e^{\frac{-2nt^2}{(b_i-a_i)^2}} \end{aligned}$$

Problem 3

Solution

(a) First for any $k \geq 0, a \geq 0$, we can have

$$P(X-\mu \geq a) = P(X-\mu+k \geq a+k)$$
$$From\ the\ def\ of\ Probability\ then\ we\ can\ have\ \leq P((X-\mu+k)^2 \geq (a+k)^2)$$

Then we can use Markov's ineuqality to get the following expression for any $k \geq 0, a \geq 0$,

$$\begin{aligned}
P((X-\mu+k)^2 \geq (a+k)^2) &= P(\left|(X-\mu+k)^2\right| \geq (a+k)^2) \\
&\leq \frac{E[\left|(X-\mu+k)^2\right|]}{(a+k)^2} \\
&\leq \frac{E[X^2+\mu^2+k^2-2X\mu+2Xk-2\mu k]}{(a+k)^2} \\
&\leq \frac{E[X^2]+E[\mu^2]+E[k^2]-2E[X\mu]-2E[\mu k]+2E[Xk]}{(a+k)^2} \\
&\leq \frac{E[X^2]+\mu^2+k^2-2\mu\mu-2\mu k+2\mu k}{(a+k)^2} \\
&\leq \frac{E[X^2]-E[X]^2+k^2}{(a+k)^2} \\
&\leq \frac{\sigma^2+k^2}{(a+k)^2}
\end{aligned}$$

So let us denote $\frac{\sigma^2+k^2}{(a+k)^2} = g(k)$, and then what we need to do is to find the minimal value of it.
Therefore $g'(k) = \frac{2(ak-\sigma^2)}{(a+k)^3}$, and because $g^2(k) >= 0$, then let $g'(k) = 0$, we can get the minimum.
Then $k = \frac{\sigma^2}{a}$, at the meanwhile $g(k) = \frac{a^2+k^2}{a^2+k^2+2\sigma^2}$.
Because $k >= 0$, and from what we have know about $\frac{a}{b} <= \frac{a+1}{b+1}$ for $a, b >= 1$.
Then let $k = 0$, we can get $g(0) = \frac{a^2}{a^2+2\sigma^2}$.
Therefore $P(X-\mu \geq a) <= \frac{a^2}{a^2+2\sigma^2} <= \frac{a^2}{a^2+\sigma^2}$.
Finally we get it.

Problem 4

Solution

(a) First from the question, we have prior distribution denoted as

$$f_\Theta(\theta) = \frac{1}{\sqrt{2\pi}\sigma_0} e^{\frac{-(\theta - x_0)^2}{2\sigma_0^2}}$$

and we also have the collection of data denoted as

$$f_{X|\Theta}(x) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma_i} e^{\frac{-(x_i - \theta)^2}{2\sigma_i^2}}$$

Then $x_1, x_2, \cdots, x_n \in R$(Observation), $x_i \in R$ and $x = (x_1, \cdots, x_n)$.
Next, we can use Bayes' rule to get the posterior PDF of Θ denoted as $f_{\Theta|X}(\theta)$.

$$f_{\Theta|X}(\theta) = \frac{f_{X|\Theta}(x) f_\Theta(\theta)}{f_X(x)}$$

But $f_X(x)$ actually is not related with parameter θ, because we need to use LOTP rule to integrate it to get $f_X(x)$.
Then we can denote it as a constant c, and get the following expression,

$$\begin{aligned} f_{\Theta|X}(\theta) &= c f_{X|\Theta}(x) f_\Theta(\theta) \\ &= c * \frac{1}{\sqrt{2\pi}\sigma_0} e^{\frac{-(\theta - x_0)^2}{2\sigma_0^2}} * \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma_i} e^{\frac{-(x_i - \theta)^2}{2\sigma_i^2}} \\ &= c * \frac{1}{(\sqrt{2\pi})^{n+1}} * \frac{1}{\prod_{i=0}^{n} \sigma_i} * e^{-\sum_{j=0}^{n} \frac{(\theta - x_j)^2}{2\sigma_j^2}} \end{aligned}$$

Then we can simplify it by denoting $c * \frac{1}{(\sqrt{2\pi})^{n+1}} * \frac{1}{\prod_{i=0}^{n} \sigma_i}$ as C.
Then we take the log of both sides and get the following expression,

$$log f_{\Theta|X}(\theta) = C - \sum_{i=0}^{n} \frac{(\theta - x_i)^2}{2\sigma_i^2}$$

Next step we need to derive θ by derivation. Denote $g(\theta) = C - \sum_{i=0}^{n} \frac{(\theta - x_i)^2}{2\sigma_i^2}$, then we can get $\theta = \frac{\sum_{i=0}^{n} \frac{x_i}{\sigma_i^2}}{\sum_{j=0}^{n} \frac{1}{\sigma_j^2}}$ by letting $g'(\theta) = 0$.
Finally we can get the posterior PDF of Θ,

$$f_{\Theta|X}(\theta) = C * e^{-\sum_{i=0}^{n} \frac{(\theta - x_i)^2}{2\sigma_i^2}}$$

and the θ, C is what we have calculated. And then we can simplify the expression and get it is a normal distribution with the mean μ and the variance σ^2.
$\mu = \frac{\sum_{i=0}^{n} \frac{x_i}{\sigma_i^2}}{\sum_{j=0}^{n} \frac{1}{\sigma_j^2}}$ and $\frac{1}{\sigma^2} = \frac{1}{\prod_{i=0}^{n} \sigma_i^2}$.

Problem 5

Solution

(a) First n independent Exponential distribution $X_1, X_2, \cdots X_n \sim Expo(\theta)$, $f_{X_i}(x) = \theta e^{-\theta x}$, $x \in (0, +\infty)$.
Then $x_1, x_2, \cdots, x_n \in R$(Observation), $x_i \in (0, +\infty)$.
And then,

$$\begin{aligned} f_X(x;\theta) &= \prod_{i=1}^{n} f_{X_i}(x_i;\theta) \\ &= \prod_{i=1}^{n} \theta e^{-\theta x_i} \\ &= \theta^n e^{-\theta(x_1+x_2+\cdots+x_n)} \end{aligned}$$

Denote $x_1 + x_2 + \cdots + x_n = S_n$, and then we can have

$$log f_X(x;\theta) = log\theta^n e^{-\theta S_n} = f(\theta)$$

Then we can get $f'(\theta) = \frac{n}{\theta} - S_n$, and $f^2(\theta) = -\frac{n}{\theta^2} \leq 0$.
Therefore $f'(\theta) = 0$, then we can get the max value of $f(\theta)$ and at the meanwhile $\theta = \frac{n}{S_n}$.
Then $\overline{\theta_{MLE}} = argmax_\theta f(\theta) = \frac{n}{S_n} = \frac{n}{x_1+x_2+\cdots+x_n}$.

(b) First n independent normal distribution $X_1, X_2, \cdots X_n \sim N(\mu, v)$, $f_{X_i}(x) = \frac{1}{\sqrt{2\pi v}} e^{-\frac{(x-\mu)^2}{2v}}$, $x \in R$.
Then $x_1, x_2, \cdots, x_n \in R$(Observation), $x_i \in R$.
And then,

$$\begin{aligned} f_X(x;\mu,v) &= \prod_{i=1}^{n} f_{X_i}(x_i;\mu,v) \\ &= \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi v}} e^{-\frac{(x-\mu)^2}{2v}} \\ &= (2\pi v)^{-\frac{n}{2}} e^{-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2v}} \end{aligned}$$

Then we can have,

$$\begin{aligned} log f_X(x;\theta) &= log((2\pi v)^{-\frac{n}{2}} e^{-\frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2v}}) \\ &= -\frac{n}{2} log(2\pi) - \frac{n}{2} log(v) - \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2v} \end{aligned}$$

Then we need divide it into two steps, one is to take the partial derivative with respect to μ, and another is v.
Then denote $g(v) = -\frac{n}{2}log(2\pi) - \frac{n}{2}log(v) - \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2v} = h(\mu)$.
And then let $g'(v) = 0 = \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{2v^2} - \frac{n}{2v}$ and $h'(\mu) = 0 = \frac{\sum_{i=1}^{n}(x_i-\mu)}{v}$,
finally we can get $v = \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{n}$, $\mu = \frac{\sum_{i=1}^{n} x_i}{n}$.
In the end $\overline{v_{MLE}} = argmax_v g(v) = \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{n}$,
and $\overline{\mu_{MLE}} = argmax_\mu h(\mu) = \frac{\sum_{i=1}^{n} x_i}{n}$.
Therefore $\overline{\theta} = (\frac{\sum_{i=1}^{n} x_i}{n}, \frac{\sum_{i=1}^{n}(x_i-\mu)^2}{n})$.