Probability & Statistics for EECS: Homework 13 # Solution

Problem 1

Let $X_1, X_2, \ldots$ be i.i.d. Expo(1).

(a) Let $N = \min\{n : X_n \geq 1\}$ be the index of the first X_j to exceed 1. Find the distribution of $N-1$ (give the name and parameters), and hence find $E(N)$.

(b) Let $M = \min\{n : X_1 + X_2 + \cdots + X_n \geq 10\}$ be the number of X_j's we observe until their sum exceeds 10 for the first time. Find the distribution of $M-1$ (give the name and parameters), and hence find $E(M)$.

(c) Let $\bar{X}_n = (X_1 + \cdots + X_n)/n$. Find the exact distribution of $\bar{X}_n$ (give the name and parameters), as well as the approximate distribution of $\bar{X}_n$ for n large (give the name and parameters).

Solution

(a) Each X_j has probability $1/e$ of exceeding 1 , so $N-1 \sim \operatorname{Geom}(1/e)$ and $E(N) = e$.

(b) Interpret $X_1, X_2, \ldots$ as the interarrival times in a Poisson process of rate 1 . Then $X_1 + X_2 + \cdots + X_j$ is the time of the j th arrival, so $M-1$ is the number of arrivals in the time interval $[0, 10)$. Thus, $M-1 \sim \operatorname{Pois}(10)$ and $E(M) = 10 + 1 = 11$.

(c) We have $X_j/n \sim \operatorname{Expo}(n)$, so $\bar{X}_n \sim \operatorname{Gamma}(n, n)$. In particular, $E\left(\bar{X}_n\right) = 1, \operatorname{Var}\left(\bar{X}_n\right) = 1/n$. By the CLT, the distribution of $\bar{X}_n$ is approximately $\mathcal{N}(1, 1/n)$ for n large.

Problem 2

Let the random variables $X_1, X_2, \ldots, X_n$ be independent with $E(X_i) = \mu$, $a \leq X_i \leq b$ for each $i = 1, \ldots, n$, where a, b are constants. Then for any $\epsilon \geq 0$, show the Hoeffding Bound holds:

$$P\left(\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| \geq \epsilon\right) \leq 2\exp\left(-\frac{2n\epsilon^2}{(b-a)^2}\right).$$

Hint: Hoeffding Lemma + Chernoff Inequality.

Solution

Let $S_n = \sum_{i=1}^{n} X_i$, According to Chernoff Tech(Inequality), for $t > 0$,

$$\begin{aligned} P(S_n - E(S_n) \geq \varepsilon') &\leq \frac{E(e^{t(S_n - E(S_n)})}{e^{t\varepsilon'}} \\ &= \frac{E(e^{t(\sum_{i=1}^{n}(X_i - E(X_i)))})}{e^{t\varepsilon'}} \\ &= \frac{\prod_{i=1}^{n} E(e^{t(X_i - E(X_i))})}{e^{t\varepsilon'}} \end{aligned}$$

Let $\varepsilon' = n\varepsilon$, we have:

$$P((\frac{1}{n}X_i - \mu) \geq \varepsilon) \leq \frac{\prod_{i=1}^{n} E(e^{t(X_i - E(X_i))})}{e^{tn\varepsilon}}$$

Because $E(X_i - E(X_i)) = 0$, $X_i \in [a, b]$, according to Hoeffding Lemma,

$$\begin{aligned} \frac{\prod_{i=1}^{n} E(e^{t(X_i - E(X_i))})}{e^{tn\varepsilon}} &\leq \frac{\prod_{i=1}^{n} e^{\frac{t^2(b-a)^2}{8}}}{e^{tn\varepsilon}} \\ &= \frac{e^{\frac{n(b-a)^2}{8}t^2}}{e^{tn\varepsilon}} \\ &= e^{\frac{n(b-a)^2}{8}t^2 - n\varepsilon t} \\ &\leq e^{-\frac{2n\varepsilon^2}{(b-a)^2}} \end{aligned}$$

Therefore,

$$\begin{aligned} P(|\frac{1}{n}X_i - \mu| \geq \varepsilon) &\leq 2P(\frac{1}{n}X_i - \mu \geq \varepsilon) \\ &\leq 2e^{-\frac{2n\varepsilon^2}{(b-a)^2}} \end{aligned}$$

Problem 3

Given a random variable X with expectation μ and variance σ^2. For any $a \geq 0$, show the following inequality holds:

$$P\left(X-\mu \geq a\right) \leq \frac{\sigma^2}{\sigma^2+a^2}.$$

Solution

Let X be a real-valued random variable with finite variance σ^2 and expectation μ, and define $Y = X - \mathbb{E}[X]$ (so that $\mathbb{E}[Y] = 0$ and $\operatorname{Var}(Y) = \sigma^2$) Then, for any $u \geq 0$, we have

$$\Pr(X-\mathbb{E}[X] \geq a) = \Pr(Y \geq a) = \Pr(Y+u \geq a+u) \leq \Pr\left((Y+u)^2 \geq (a+u)^2\right) \leq \frac{\mathbb{E}\left[(Y+u)^2\right]}{(a+u)^2} = \frac{\sigma^2+u^2}{(a+u)^2}$$

the last inequality being a consequence of Markov's inequality. As the above holds for any choice of $u \in \mathbb{R}$, we can choose to apply it with the value that minimizes the function $u \geq 0 \mapsto \frac{\sigma^2+u^2}{(a+u)^2}$. By differentiating, this can be seen to be $u_* = \frac{\sigma^2}{a}$, leading to

$$\Pr(X-\mathbb{E}[X] \geq a) \leq \frac{\sigma^2+u_*^2}{\left(a+u_*\right)^2} = \frac{\sigma^2}{a^2+\sigma^2} \text{ if } a>0$$

Problem 4

We observe a collection $X = (X_1, \ldots, X_n)$ of random variables, with an unknown common mean whose value we wish to infer. We assume that given the value of the common mean, the X_i are normal and independent, with known variances $\sigma_1^2, \ldots, \sigma_n^2$. We model the common mean as a random variable Θ, with a given normal prior (known mean x_0 and known variance σ_0^2). Find the posterior PDF of Θ.

Solution

According to the problem, $X_i|\Theta = \theta \sim \mathcal{N}(\theta, \sigma_i^2)$, and $\Theta \sim \mathcal{N}(x_0, \sigma^2)$, therefore by Bayesian inference:

$$
\begin{aligned}
f_{\Theta|X}(\theta \mid x) &\propto f_{X|\Theta}(x \mid \theta) \cdot f_\Theta(\theta) \\
&= \prod_{i=1}^{n} \left(\frac{1}{\sqrt{2\pi}\sigma_i} \exp\left(-\frac{(x_i - \theta)^2}{2\sigma_i^2} \right) \right) \cdot \frac{1}{\sqrt{2\pi}\sigma_0} \exp\left(-\frac{(\theta - x_0)^2}{2\sigma_0^2} \right) \Bigg) \\
&\propto \exp\left(-\sum_{i=0}^{n} \frac{(\theta - x_i)^2}{2\sigma_i^2} \right)
\end{aligned}
$$

Suppose $\Theta|X = (x_1, \ldots, x_n) \sim \mathcal{N}(x, \sigma^2)$, so:

$$
f_{\Theta|X}(\theta|x) \propto exp(-\frac{(\theta - x)^2}{2\sigma^2}) = \exp\left(-\sum_{i=0}^{n} \frac{(\theta - x_i)^2}{2\sigma_i^2} \right)
$$

We can solve the x and σ:

$$
\begin{aligned}
x &= \frac{\sum_{i=0}^{n} \frac{x_i}{\sigma_i^2}}{\sum_{i=0}^{n} \frac{1}{\sigma_i^2}} \\
\sigma &= (\frac{1}{\sum_{i=0}^{n} \frac{1}{\sigma_i^2}})^{1/2}
\end{aligned}
$$

Therefore, $X \mid \Theta \sim N\left(\sum_{i=0}^{n} \frac{x_i}{\sigma_i^2} / \sum_{i=0}^{n} \frac{1}{\sigma_i^2}, 1/\sum_{i=0}^{n} \frac{1}{\sigma_i^2} \right)$.

Problem 5

(a) We wish to estimate the parameter for an exponential distribution, denoted by θ, based on the observations of n independent random variables $X_1, \ldots, X_n$, where $X_i \sim \text{Expo}(\theta)$. Find the MLE of θ.

(b) We wish to estimate the mean μ and variance ν of a normal distribution using n independent observations $X_1, \ldots, X_n$, where $X_i \sim \mathcal{N}(\mu, \nu)$. Find the MLE of the parameter vector $\theta = (\mu, \nu)$.

Solution

(a) By MLE, we need to solve $\theta = \text{argmax}_\theta\ P(X|\theta)$,

$$P(X|\theta) = \prod_{i=1}^{n} \theta e^{-\theta x_i} = \theta^n e^{-\theta \sum_{i=1}^{n} x_i}$$

By solving $\theta = \text{argmax}_\theta\ P(X|\theta) = \text{argmax}_\theta\ \theta^n e^{-\theta \sum_{i=1}^{n} x_i}$, we have the MLE of θ:

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

(b) By MLE, we need to solve $\theta = \text{argmax}_\theta\ P(X|\theta)$,

$$P(X|\theta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\nu}} exp(-\frac{(x_i - \mu)^2}{2\nu}) = \frac{1}{\sqrt{2\pi\nu}} exp(-\sum_{i=1}^{n} \frac{(x_i - \mu)^2}{2\nu})$$

By solving $\theta = \text{argmax}_\theta\ P(X|\theta) = \text{argmax}_\theta\ \frac{1}{\sqrt{2\pi\nu}} exp(-\sum_{i=1}^{n} \frac{(x_i-\mu)^2}{2\nu})$, we have the MLE of $\theta = (\mu, \nu)$:

$$\mu = \frac{1}{n} \sum_{i=1}^{n} x_i$$
$$\nu = \frac{1}{n} \sum_{i=1}^{n} (x_i - \mu)^2$$