Solutions for Homework 7

(1) Since $f_X(x) = (1/\beta)e^{-x/\beta}$,

$$\begin{aligned} \mathrm{E}(X) &= \int x\beta^{-1}e^{-x/\beta}dx = \beta \\ \mathrm{E}(X^2) &= \int x^2\beta^{-1}e^{-x/\beta}dx = 2\beta^2 \\ \mathrm{Var}(X) &= 2\beta^2 - \beta^2 = \beta^2. \end{aligned}$$

$$\begin{aligned} P(|X-\mu_X| \geq k\sigma_X) &= P(|X-\beta| \geq k\beta) \\ &= P(X-\beta \geq k\beta) = P(X \geq (k+1)\beta) \\ &= \int_{k(1+\beta)}^{\infty} f_X(x)\,dx \\ &= e^{-k(1+\beta)/\beta}. \end{aligned}$$

Chebyshev's inequality gives,

$$P(|X-\mu_X| \geq k\sigma_X) \leq \frac{\sigma_X^2}{k^2\sigma_X^2} = \frac{1}{k^2}.$$

Comment: Using elementary calculus, one can show that $k - 2\log k > 0$ for all $k > 0$. Hence, for any $\beta > 0$, $k(1+(1/\beta)) \geq k > 2\log k$. Exponentiating we see that $e^{-k(1+\beta)/\beta} \leq 1/k^2$ which confirms that Chebyshev only gives an upper bound.

(2) $E(X) = V(X) = \lambda$. So,

$$\begin{aligned} \mathrm{P}\left(X \geq 2\lambda\right) &= \mathrm{P}\left(X-\lambda \geq \lambda\right) \\ &\leq \mathrm{P}\left(|X-\lambda| \geq \lambda\right) \\ &\leq \frac{\lambda}{\lambda^2} = \frac{1}{\lambda}. \end{aligned}$$

(3) First, recall that $\mathrm{E}(\overline{X}_n) = p$ and $\mathrm{Var}(\overline{X}_n) = \mathrm{Var}(\mathrm{X}_1)/n = p(1-p)/n$. From Chebyshev,

$$\begin{aligned} \mathrm{P}\left(|\overline{X}_n - p| > \epsilon\right) &\leq \frac{\mathrm{Var}(\overline{X}_n)}{\epsilon^2} \\ &= \frac{p(1-p)}{n\epsilon^2}. \end{aligned}$$

From Hoeffding,

$$\mathrm{P}\left(|\overline{X}_n - p| > \epsilon\right) \leq 2e^{-2n\epsilon^2}.$$

Because of the exponential term, Hoeffding gives a smaller bound.

Comment: To make this observation precise, note that, for any $c > 0$, $e^x > x + c$ for all large $x > 0$. To see this, recall that $e^x = 1 + x + x^2/2 + x^3/3! + \cdots$. Now $1 + x^2/2 + x^3/3! + \cdots > c$ for all large x. Hence, $e^x = 1 + x + x^2/2 + x^3/3! + \cdots > c + x$ for all large x. Take $x = 2n\epsilon^2$ and $c = \log 1/(p(1-p))$ and conclude that $2n\epsilon^2 > \log(2n\epsilon^2) + \log 1/(p(1-p))$ for all large n. Exponentiate both sides and conclude that $2e^{-2n\epsilon^2} < p(1-p)/(n\epsilon^2)$ for all large n.

(4a) Let us write

$$S_n^2 = \frac{\sum_i (X_i \overline{X}_n)^2}{n-1} = \frac{\sum_i X_i^2 - n\overline{X}_n^2}{n-1}.$$

Now, $E\sum_i X_i^2 = nE(X_1)^2$. To compute $\mathrm{E}\left(\overline{X}_n^2\right)$ we will make use of the following fact. If $a_1, \ldots, a_n$ are real numbers then $(\sum_i a_i)^2 = \sum_i a_i^2 + \sum_{i \neq j} a_i a_j$ and the second sum has $n^2 - n = n(n-1)$ terms. So,

$$\begin{aligned}
\mathrm{E}\left(\overline{X}_n^2\right) &= \mathrm{E}\left(\frac{1}{n}\sum_{i=1}^n X_i\right)^2 \\
&= \frac{1}{n^2}\mathrm{E}\left(\sum_{i=1}^n X_i\right)^2 \\
&= \frac{1}{n^2}\mathrm{E}\left(\sum_{i=1}^n X_i^2 + \sum_{i \neq j} X_i X_j\right) \\
&= \frac{1}{n^2}\left(\sum_{i=1}^n \mathrm{E}(X_i)^2 + \sum_{i \neq j} E(X_i X_j)\right) \\
&= \frac{1}{n^2}\left(\sum_{i=1}^n \mathrm{E}(X_i)^2 + \sum_{i \neq j} E(X_i)E(X_j)\right) \\
&= \frac{1}{n^2}\left(\sum_{i=1}^n \mathrm{E}(X_1)^2 + \sum_{i \neq j} \mu^2\right) \\
&= \frac{1}{n^2}\left(nE(X_1)^2 + n(n-1)\mu^2\right).
\end{aligned}$$

Thus,

$$E(S_n^2) = \frac{nE(X_1^2) - n\frac{\left(nE(X_1)^2 + n(n-1)\mu^2\right)}{n^2}}{n-1} = E(X_1^2) - \mu^2 = \sigma^2.$$

(4b) We can write

$$S_n^2 = \left(\frac{n}{n-1}\right)\frac{1}{n}\sum_{i=1}^{n} X_i^2 - \left(\frac{n}{n-1}\right)\overline{X}_n^2.$$

By the law of large numbers,

$$\frac{1}{n}\sum_{i=1}^{n} X_i^2 \xrightarrow{p} E(X_1^2).$$

Since $(n-1)/n \to 1$, we also have that

$$\left(\frac{n}{n-1}\right)\frac{1}{n}\sum_{i=1}^{n} X_i^2 \xrightarrow{p} E(X_1^2).$$

Also, by the law of large numbers, $\overline{X}_n \xrightarrow{p} \mu$. Since $g(y) = y^2$ is a continuous function, $\overline{X}_n^2 \xrightarrow{p} \mu^2$. So,

$$\left(\frac{n}{n-1}\right)\frac{1}{n}\sum_{i=1}^{n} X_i^2 - \left(\frac{n}{n-1}\right)\overline{X}_n^2 \xrightarrow{p} E(X_1^2) - \mu^2 = \sigma^2.$$

(5) Let $\mu_n = E(X_n)$. Then,

$$\begin{aligned}
\mathrm{E}(X_n - b)^2 &= \mathrm{E}(X_n - \mu_n + \mu_n - b)^2 \\
&= \mathrm{E}\left[(X_n - \mu_n)^2 + (\mu_n - b)^2 + 2(X_n - \mu_n)(\mu_n - b)\right] \\
&= \mathrm{E}(X_n - \mu_n)^2 + (\mu_n - b)^2 + 2(\mu_n - b)\mathrm{E}(X_n - \mu_n) \\
&= \mathrm{E}(X_n - \mu_n)^2 + (\mu_n - b)^2 \\
&= \mathrm{Var}(X_n) + (\mu_n - b)^2.
\end{aligned}$$

From this last expression we see that if $\mu_n \to b$ and $\mathrm{Var}(X_n) \to 0$ then $\mathrm{E}(X_n - b)^2 \to 0$. Conversely, if $\mathrm{E}(X_n - b)^2 \to 0$ then $\mathrm{Var}(X_n) + (\mu_n - b)^2 \to 0$. Hence, $\mathrm{Var}(X_n) \to 0$ and $(\mu_n - b)^2 \to 0$.

(6) $E(\overline{X}_n) = \mu$ and $\text{Var}(\overline{X}_n) = \sigma^2/n$. So $E(\overline{X}_n - \mu)^2 = \text{Var}(\overline{X}_n) = \sigma^2/n \to 0$. Hence, $\overline{X}_n \stackrel{q.m.}{\to} \mu$.

(7) Fix $\epsilon > 0$. Then, for all $n > 1/\epsilon$, $P(|X_n| > \epsilon) = P(X_n = n) = 1/n^2$. Thus, $P(|X_n| > \epsilon) \to 0$ and so $X_n \stackrel{p}{\to} 0$. But $E((X_n - 0)^2) = E(X_n^2) = (1/n^2) \times (1 - (1/n^2)) + (n^2) \times (1/n^2) = (1/n^2) - (1/n^4) + 1 \not\to 0$ so X_n does not converge in quadratic mean.

(8a) If $p \notin C_n$ then $|p - \hat{p}_n| > \epsilon_n$. Hence,

$$\begin{aligned} \text{P}(p \notin C_n) &= \text{P}(|p - \hat{p}_n| > \epsilon_n) \\ &\leq 2e^{-2n\epsilon^2} \quad \text{Hoeffding's inequality} \\ &= \alpha \end{aligned}$$

where the last line follows from plugging in the definition of ϵ_n.

(8b) The interval tends to over-cover.

(8c) When $n = 74$, the length $2\epsilon_n \leq .05$.