Solution for Assignment 15

Fan Zeteng

December 11, 2022

PROBLEM 1. Suppose that a discrete random variable X has finite $k-th$ moment, i.e., $E(|X|^k) < \infty, (k > 0$, but k may not be a positive integer). Show that for any $\epsilon > 0$,

$$P\{|X| \geq \epsilon\} \leq \frac{E(|X|^k)}{\epsilon^k}.$$

SOLUTION. For any $\epsilon > 0$,

$$\begin{aligned} P\{|X| \geq \epsilon\} =& \sum_{k=1}^{\infty} P(X = x_k) 1_{|x_k|>\epsilon} \\ \leq& \sum_{k=1}^{\infty} P(X = x_k) 1_{|x_k|>\epsilon} [\frac{|x_k|}{\epsilon}]^k \\ \leq& \frac{1}{\epsilon^k} \sum_{k=1}^{\infty} P(X = x_k) |x_k|^k \\ =& \frac{E(|X|^k)}{\epsilon^k}. \end{aligned}$$

PROBLEM 2. Suppose that $\{X_1, X_2, \cdots, X_n, \cdots\}$ is a sequence of independent r.v.s (not necessarily with the same distribution), each with finite (but not necessarily with the same)mean and uniformly bounded variance by $M < \infty, (i.e., Var(X_i) \leq M, \forall i > 1)$. Let $\overline{X} = \frac{1}{n} \sum_{i=1} nX_i$ be the sample mean. Show that for any $\epsilon > 0$, we have

$$\lim_{n\to\infty} P\{|\overline{X}_n - E\overline{X}_n| > \epsilon\} = 0.$$

Solution. Given a fixed ϵ, for any $n \le 1$, we know form problem 1 that

$$\begin{aligned} P\{|\overline{X}_n - E\overline{X}_n| \ge \epsilon\} \le & \frac{E(|\overline{X}_n - E\overline{X}_n|^2)}{\epsilon^2} \\ = & \frac{Var(\overline{X}_i)}{\epsilon^2} \\ = & \frac{\sum_{i=1}^n Var(X_i)}{n^2\epsilon^2} \\ \le & \frac{M}{n\epsilon^2}. \end{aligned}$$

So

$$\overline{\lim_{n\to\infty}} P\{|\overline{X}_n - E\overline{X}_n| > \epsilon\} \le \lim_{n\to\infty} \frac{M}{n\epsilon^2} = 0.$$

Which indicates

$$\lim_{n\to\infty} P\{|\overline{X}_n - E\overline{X}_n| > \epsilon\} = 0.$$

Problem 3. Suppose that $X_1, X_2, \cdots$ is a sequence of i.i.d. r.v.s with common mean 1 and variance 16. Let n be sufficiently large and $Y = X_1 + X_2 + \cdots + X_n$. Estimate the value of $P\{2.608 < Y \le 4.4124\}$. (Hint: using the central limit theorem and normal approximation method.)

Solution. By the method of central limit theorem, when n is sufficiently large, we konw $Y = X_1 + X_2 + \cdots + X_n$ converges to $N(n, 16n)$. And we can estimate it as

$$\begin{aligned} P\{2.608 < Y \le 4.4124\} = & P\{\frac{2.608 - n}{4\sqrt{n}} < \frac{Y - n}{4\sqrt{n}} \le \frac{4.4124 - n}{4\sqrt{n}}\} \\ = & P\{\frac{2.608 - n}{4\sqrt{n}} < Z \le \frac{4.4124 - n}{4\sqrt{n}}\} \\ = & \Phi(\frac{4.4124 - n}{4\sqrt{n}}) - \Phi(\frac{2.608 - n}{4\sqrt{n}}). \end{aligned}$$

Problem 4. Let $X_{nn\ge1}$ be a sequence of i.i.d. random variables with a common normal distribution $N(20, 25)$. Let

$$Y = \sum_{i=1}^{6} X_i, W = \sum_{i=1}^{30} X_i, \overline{X}_{30} = \frac{1}{30}W.$$

(i) What distributions do Y, W and $\overline{X}_{30}$ obey?

(ii) Find the probability that the random variable $\overline{X}_{30}$ is between 19 and 21, i.e., find $P(19 < X_{30} < 21)$. Also, find the probability that W is greater than 650, i.e., find $P(W > 650)$.

(iii) Are Y and W uncorrelated or correlated? If you think they are uncorrelated, then provide a proof. If you think they are correlated, then state whether they are positively correlated or negatively correlated. Specify your reasons clearly! Also find the correlation coefficient of Y and W, i.e., $\rho(Y, W)$ to support your conclusions.

(iv) Let $\overline{X}_n = \frac{1}{n}\sum\limits_{i=1}^{n} X_i$. Find the moment generating function of $\overline{X}_n$, denoted by $M_{\overline{X}_n}(t)$. Let $\varphi(t) = \lim\limits_{n\to\infty} M_{\overline{X}_n}(t)$. Is $\varphi(t)$ the moment generating function of some normal random variable? If your answer is positive, specify this normal distribution. If your answer is negative, specify some reasons to support your conclusion, and state the random variable (a constant is also viewed as a random variable) whose moment generating function is $\varphi(t)$. Let $Y_n = \frac{\overline{X}_n - 20}{5\sqrt{n}}$. Find the moment generating function of Y_n , denoted by $M_{Y_n}(t)$. Let $\psi(t) = \lim\limits_{n\to\infty} M_{Y_n}(t)$. Answer the same questions for $\psi(t)$.

Solution.

(i) By the property of m.g.f.,we know

$$\begin{aligned}
M_Y(t) =& M_{\sum\limits_{i=1}^{6} X_i}(t) \\
=& \prod_{i=1}^{6} M_{X_i}(t) \\
=& (e^{20t+\frac{25}{2}t^2})^6 \\
=& e^{120t+\frac{6*25}{2}t^2}.
\end{aligned}$$

So from the unique property, $Y \sim N(6*20, 6*25)$.By the same way, we have

$$W \sim N(30*20, 30*25), \overline{X}_{30} \sim N(20, \frac{25}{30}).$$

(ii) Sin $\frac{\overline{X}_{30}-20}{\frac{5}{\sqrt{30}}}$ is standard normal r.v., we have

$$\begin{aligned} P(19 < X_{30} < 21) =& P(\sqrt{30}\frac{19-20}{5} < \sqrt{30}\frac{\overline{X}_{30}-20}{5} < \sqrt{30}\frac{21-20}{5}) \\ =& \Phi(\frac{\sqrt{30}}{5}) - \Phi(-\frac{\sqrt{30}}{5}). \end{aligned}$$

By the same way, $P(W > 650) = P(\frac{W-600}{5\sqrt{30}} > \frac{650-600}{5\sqrt{30}}) = 1 - \Phi(\frac{5}{\sqrt{30}}) = \Phi(-\frac{5}{\sqrt{30}})$

(iii) Intuitively, we think they are positively correlated, since $W = Y + (W - Y)$, while the second part is independent with Y. And we can compute the correlation coefficient:

$$\begin{aligned} Cov(Y, W) =& Cov(Y, Y + (W - Y)) \\ =& Cov(Y, Y) \\ =& Var(Y) = 6*25. \end{aligned}$$

$$\rho(Y, W) = \frac{Cov(Y, W)}{\sqrt{Var(Y)Var(W)}} = \frac{6*25}{sqrt30*25*6*25} = \frac{\sqrt{5}}{5}$$

(iv) By the same way as (i), we know $M_{\overline{X}_n}(t) = e^{20t+\frac{25}{2*n}t^2}$.And

$$\begin{aligned} \varphi(t) =& \lim_{n\to\infty} M_{\overline{X}_n}(t) \\ =& \lim_{n\to\infty} e^{20t+\frac{25}{2*n}t^2} \\ =& e^{20t}. \end{aligned}$$

It is just the m.g.f. of the const r.v. $X \equiv 20$.We regard it as $N(20, 0)$, a degenerate normal r.v..As for $Y_n = \frac{\overline{X}_n-20}{5\sqrt{n}}$, by the property of m.g.f., we know

$$
\begin{aligned}
M_{Y_n}(t) =& e^{-4\sqrt{n}t} M_{\overline{X}_n}(\frac{t}{5/\sqrt{n}}) \\
=& e^{-4\sqrt{n}t} e^{4\sqrt{n}t + \frac{25}{2*n}\frac{n}{25}t^2} \\
=& e^{\frac{t^2}{2}}.
\end{aligned}
$$

It is just the m.g.f. of standard normal r.v..