Solutions to Assignment #18

1. Let a denote a real number and X_a be a discrete random variable with pmf

$$p_{X_a}(x) = \begin{cases} 1 & \text{if } x = a; \\ 0 & \text{elsewhere.} \end{cases}$$

   (a) Compute the cdf for X_a and sketch its graph.

   (b) Compute the mgf for X_a and determine $E(X_a)$ and $\text{var}(X_a)$.

Solution:

(a) For $x < a$, we get that

$$F_{X_a}(x) = \Pr(X_a \leqslant x) = 0.$$

If $x \geqslant a$, then

$$F_{X_a}(x) = \Pr(X_a \leqslant x) = \Pr(X_a = a) = 1.$$

Thus,

$$F_{X_a}(x) = \begin{cases} 0 & \text{if } x < a, \\ 1 & \text{if } x \geqslant a. \end{cases}$$

The graph of F_{X_a} is pictured in Figure 1

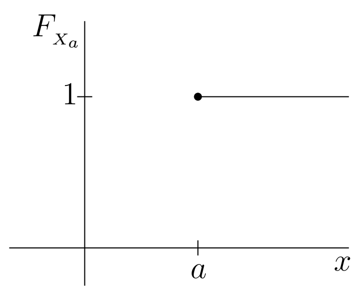

Figure 1: Cumulative Distribution Function for X_a

(b) The mgf for X_a is

$$\psi_{X_a}(t) = E(e^{tX_a}) = e^{ta}\, p_{X_a}(a) = e^{at} \quad \text{for all } t \in \mathbb{R}.$$

Differentiating with respect to t we obtain

$$\psi'_{X_a}(t) = ae^{at} \quad \text{for all } t \in \mathbb{R},$$

and

$$\psi''_{X_a}(t) = a^2 e^{at} \quad \text{for all } t \in \mathbb{R}.$$

It then follows that the expected value of X_a is

$$E(X_a) = \psi'_{X_a}(0) = a;$$

its second moment is

$$E(X_a^2) = \psi''_{X_a}(0) = a^2;$$

Thus, the variance of X_a is

$$\text{var}(X_a) = E(X_a^2) - a^2 = 0.$$

□

2. Let (X_k) denote a sequence of independent identically distributed random variables such that $X_k \sim \text{Normal}(\mu, \sigma^2)$ for every $k = 1, 2, \ldots$, and for some $\mu \in \mathbb{R}$ and $\sigma > 0$. For each $n \geqslant 1$, define

$$\overline{X}_n = \frac{X_1 + X_2 + \cdots + X_n}{n}.$$

(a) Determine the mgf, $\psi_{\overline{X}_n}(t)$, for $\overline{X}_n$, and compute $\lim\limits_{n\to\infty} \psi_{\overline{X}_n}(t)$.

(b) Find the limiting distribution of $\overline{X}_n$ as $n \to \infty$. (*Hint:* Compare your answer in part (a) to your answer in part (b) of problem 1.)

Solution:

(a) Compute

$$\begin{aligned}
\psi_{\overline{X}_n}(t) &= E(e^{t\overline{X}_n}) \\
&= \psi_{X_1+X_2+\cdots+X_n}\left(\frac{t}{n}\right) \\
&= \psi_{X_1}\left(\frac{t}{n}\right)\psi_{X_2}\left(\frac{t}{n}\right)\cdots\psi_{X_n}\left(\frac{t}{n}\right),
\end{aligned}$$

since the X_i's are linearly independent. Thus, given the X_i's are identically distributed Normal(μ, σ^2),

$$\begin{aligned} \psi_{\overline{X}_n}(t) &= \left[\psi_{X_1}\left(\frac{t}{n}\right)\right]^n \\ &= \left[e^{\mu(t/n)+\sigma^2(t/n)^2/2}\right]^n \\ &= e^{\mu t+\sigma^2 t^2/2n}. \end{aligned}$$

We then get that for $t \neq 0$,

$$\lim_{n\to\infty} \psi_{\overline{X}_n}(t) = e^{\mu t}.$$

On the other hand, if $t = 0$, $\psi_{\overline{X}_n}(0) = 1$ for all n. We therefore get that

$$\lim_{n\to\infty} \psi_{\overline{X}_n}(t) = e^{\mu t} \quad \text{for all} \quad t \in \mathbb{R}.$$

(b) By problem 1, the limit obtained in the previous part is the mgf of the discrete random variable X_μ which has pmf

$$p_{X_\mu}(x) = \begin{cases} 1 & \text{if } x = \mu; \\ 0 & \text{elsewhere.} \end{cases}$$

By the mgf Convergence Theorem, the sequence of sample means, $(\overline{X}_n)$, converges in distribution to X_μ. In other words, for every $\varepsilon > 0$,

$$\lim_{n\to\infty} \Pr(\overline{X}_n \leqslant \mu + \varepsilon) = \Pr(X_\mu) \leqslant \mu + \varepsilon) = 1,$$

while

$$\lim_{n\to\infty} \Pr(\overline{X}_n \leqslant \mu - \varepsilon) = \Pr(X_\mu) \leqslant \mu - \varepsilon) = 0.$$

It then follows that

$$\lim_{n\to\infty} \Pr(\mu - \varepsilon < \overline{X}_n \leqslant \mu + \varepsilon) = 1$$

for all $\varepsilon > 0$. Thus, with probability 1, the sample mean will within an arbitrary distance of the mean of the distribution as the sample size increases to infinity.

□

3. Let (X_k) and $\overline{X}_n$ be defined as in the previous problem. Define $Z_n = \dfrac{\overline{X}_n - \mu}{\sigma/\sqrt{n}}$ for all $n \geqslant 1$.

   (a) Determine the mgf, $\psi_{\overline{Z}_n}(t)$, for $\overline{Z}_n$, and compute $\lim\limits_{n\to\infty} \psi_{\overline{Z}_n}(t)$.

   (b) Find the limiting distribution of $\overline{Z}_n$ as $n \to \infty$.

   ***Solution**:*

   (a) Compute

$$
\begin{aligned}
\psi_{Z_n}(t) &= E(e^{tZ_n}) \\
&= E(e^{(t\sqrt{n}/\sigma)\overline{X}_n}\ e^{-\mu t\sqrt{n}/\sigma}) \\
&= e^{-\mu t\sqrt{n}/\sigma}\ \psi_{\overline{X}_n}\left(\frac{t\sqrt{n}}{\sigma}\right) \\
&= e^{-\mu t\sqrt{n}/\sigma}\quad e^{\mu(t\sqrt{n}/\sigma)+\sigma^2(t\sqrt{n}/\sigma)^2/2n},
\end{aligned}
$$

   where we have used the expression for the mgf of $\overline{X}_n$ computed in the previous problem. It then follows that

$$
\psi_{Z_n}(t) = e^{\sigma^2(t\sqrt{n}/\sigma)^2/2n} = e^{t^2/2}
$$

   for all $n = 1, 2, 3, \ldots$ It then follows that

$$
\lim_{n\to\infty} \psi_{Z_n}(t) = e^{t^2/2} \quad \text{for all} \quad t \in \mathbb{R}.
$$

   (b) Note that the limit obtained in the previous part is the mgf of the standard normal random variable $Z \sim \text{Normal}(0, 1)$. It then follows, by the mgf Convergence Theorem that Z_n converges in distribution to $Z \sim \text{Normal}(0, 1)$ as $n \to \infty$.

□

4. Let (Y_n) be a sequence of discrete random variables having pmfs

$$p_{Y_n}(y) = \begin{cases} 1 & \text{if } y = n, \\ 0 & \text{elsewhere.} \end{cases}$$

Compute the mgf of Y_n for each $n = 1, 2, 3, \ldots$

Does $\lim\limits_{n\to\infty} \psi_{Y_n}(t)$ exist for any t in an open interval around 0?

Does the sequence (Y_n) have a limiting distribution? Justify your answer.

***Solution**:* Compute $\psi_{Y_n}(t) = E(e^{tY_n}) = e^{tn}p_{Y_n}(n) = e^{nt}$, for all $t \in \mathbb{R}$.

Observe that for $t > 0$, $\psi_{Y_n}(t) \to \infty$ as $n \to \infty$. Therefore, $\lim\limits_{n\to\infty} \psi_{Y_n}(t)$ does not exist for t in an open interval around 0.

Now, for any $y \in \mathbb{R}$, here exists a natural number n_o such that $n \geqslant n_o$ implies that $x < n$. Consequently, for all $n \geqslant n_o$,

$$\Pr(Y_n \leqslant y) = 0.$$

where therefore conclude that

$$\lim_{n\to\infty} \Pr(Y_n \leqslant y) = 0, \quad \text{for all } y \in \mathbb{R}.$$

However, 0 cannot be a cdf for any random variable. Hence, (Y_n) does not have a limiting distribution. □

5. Let $q = 0.95$ denote the probability that a person, in certain age group, lives at least 5 years.

   (a) If we observe 60 people from that group and assume independence, what is the probability that at least 56 of them live 5 years or more?

   (b) Find and approximation to the result of part (a) using the Poisson distribution.

***Solution**:*

(a) Let X denote the number of people from the group that will live 5 years or more. Then, $X \sim \text{Binomial}(q, 60)$. Consequently, the probability that at least 56 of them live 5 years or more is

$$\Pr(X \geqslant 56) = \sum_{k=56}^{60} \binom{60}{k} q^k (1-q)^{60-k}.$$

Note that we can also compute this probability as

$$\Pr(X \geqslant 56) = 1 - \Pr(X \leqslant 55) = 1 - F_X(55).$$

Using MS-Excel to make this calculation we get that

$$\Pr(X \geqslant 56) \approx 0.82 \text{ or } 82\%.$$

(b) The Poisson approximation to the a Binomial(p, n) is appropriate when p is small and n is large. Thus, instead of looking at the probability, q, of living 5 or more years, we look at the complementary probability $p = 1-q = 0.05$ of living less than five years. If we let Y denote the number of people from the group that will live less than five years, then $Y \sim \text{Binomial}(p, 60)$. The event $(X \geqslant 56)$ is then equivalent to $(Y \leqslant 4)$, and so we are interested in approximating

$$\Pr(Y \leqslant 4)$$

by a Poisson(λ) distribution with $\lambda = np = 3$. We then get that

$$\begin{aligned} \Pr(Y \leqslant 4) &\approx \sum_{k=0}^{4} \frac{3^k}{k!}\, e^{-3} \\ &= e^{-3} + 3e^{-3} + \frac{9}{2}e^{-3} + \frac{27}{6}e^{-3} + \frac{81}{24}e^{-3} \\ &\approx 0.82 \text{ or } 82\%. \end{aligned}$$

Thus, $\Pr(X \geqslant 56) \approx 0.82$ or 82% as in part (a).

□