Tutorial T12

Example T12.1: Let $X \sim \mathsf{Exp}(\lambda)$, $Y \sim \mathsf{Exp}(\lambda)$ and X and Y independent. What is the pdf of $Z = X + Y$?

Solution:

$$\begin{aligned}
M_Z(t) &= M_{X+Y}(t) \\
&= E\left[e^{t(X+Y)}\right] \\
&= E\left[e^{tX}e^{tY}\right] \\
&= E\left[e^{tX}\right]E\left[e^{tY}\right] && \text{[since X and Y are independent]} \\
&= M_X(t)M_Y(t) \\
&= \frac{\lambda}{\lambda - t}\frac{\lambda}{\lambda - t} \\
&= \frac{\lambda^2}{(\lambda - t)^2}
\end{aligned}$$

This is the same MGF as we saw in Tutorial Example T11.3. So the pdf of Z is the pdf from there:

$$f_Z(z) = \begin{cases} \lambda^2 z e^{-\lambda z} & z \geq 0 \\ 0 & \text{else.} \end{cases}$$

We can use this to verify $E[Z]$ and $E[Z^2]$ from there too:

$$\begin{aligned}
E[Z] &= E[X + Y] = E[X] + E[Y] = \frac{1}{\lambda} + \frac{1}{\lambda} = \frac{2}{\lambda} \\
E[Z^2] &= E[(X + Y)^2] \\
&= E[X^2] + 2E[X]E[Y] + E[Y^2] \\
&= \frac{2}{\lambda^2} + 2\frac{1}{\lambda}\frac{1}{\lambda} + \frac{2}{\lambda^2}
\end{aligned}$$

$$= \frac{6}{\lambda^2}$$

Example T12.2: Let $X \sim U(0,1)$.

a) Compute $P[X \geq 1/4]$, $P[X \geq 1/2]$, $P[X \geq 3/4]$ and $P[X \geq 1]$ exactly.

b) Use Markov's inequality to find upper bounds to these.

Solution:

$$f_X(x) = \begin{cases} 1 & 0 < x < 1 \\ 0 & \text{else} \end{cases}$$

a)

$$\begin{aligned} P[X \geq 1/4] &= \int_{1/4}^{\infty} f_X(x)dx = \int_{1/4}^{1} 1 \; dx = 3/4 \\ P[X \geq 1/2] &= 1/2 \\ P[X \geq 3/4] &= 1/4 \\ P[X \geq 1] &= 0 \end{aligned}$$

b) Markov inequality: If X is non-negative and $a > 0$, then $P[X \geq a] \leq E[X]/a$.

$$\begin{aligned} P[X \geq 1/4] &\leq E[X]/(1/4) = (1/2)/(1/4) = 2 \\ P[X \geq 1/2] &\leq (1/2)/(1/2) = 1 \\ P[X \geq 3/4] &\leq (1/2)/(3/4) = 2/3 \\ P[X \geq 1] &\leq (1/2)/(1) = 1/2 \end{aligned}$$

Example T12.3: Let $X \sim U(-1,1)$.

a) Compute $P[|X| \geq 1/4]$, $P[|X| \geq 1/2]$, $P[|X| \geq 3/4]$ and $P[|X| \geq 1]$ exactly.

b) Use Chebyshev's inequality to find upper bounds to these.

Solution:

a)

$$
\begin{aligned}
P[|X| \geq 1/4] &= P[\{X \leq -1/4\} \cup \{X \geq 1/4\}] = 3/8 + 3/8 = 3/4 \\
P[|X| \geq 2/4] &= 2/4 \\
P[|X| \geq 3/4] &= 1/4 \\
P[|X| \geq 1] &= 0
\end{aligned}
$$

b) Chebyshev: If X has mean μ and variance σ^2 and $k > 0$ then $P[|X - \mu| \geq k] \leq \sigma^2/k^2$.

$$E[X] = 0, \qquad Var[X] = 1/3$$

$$
\begin{aligned}
P[|X| \geq 1/4] = P[|X - 0| \geq 1/4] &\leq \frac{1/3}{(1/4)^2} = \frac{16}{3} \\
P[|X| \geq 2/4] &\leq \frac{1/3}{(2/4)^2} = \frac{4}{3} \\
P[|X| \geq 3/4] &\leq \frac{1/3}{(3/4)^2} = \frac{16}{27} \\
P[|X| \geq 4/4] &\leq \frac{1/3}{(4/4)^2} = \frac{1}{3}
\end{aligned}
$$

Example T12.4: The lifetime of a certain type of HDD is known to be never greater than 7 years.

We want to estimate the average lifetime μ of this HDD type by computing the numerical average Y of the measured lifetimes X_n of n HDD:

$$Y = \frac{X_1 + \cdots + X_n}{n}$$

Assume lifetimes are iid.

a) Find an upper bound to $Var[X_1]$.

b) How large should n be if we want $SD[Y] \leq 0.1$ years?

c) How large should n be if we want Chebyshev to guarantee an estimate that is within 0.5 years with prob. 0.99?

d) Assuming that n is large enough that the CLT applies, repeat c) using the CLT.

Solution: a)

$$Var[X_1] = E[X_1^2] - (E[X_1])^2 \leq E[X_1^2] \leq 7^2 \text{ years}^2$$

b) $SD[Y] = 0.1$ years $\Leftrightarrow Var[Y] = 0.01$ years^2.

$$Var[Y] = \frac{Var[X_1]}{n} \leq \frac{49 \text{ years}^2}{n}$$

So $n = 4900$ is guaranteed to give $SD[Y] \leq 0.1$ years.

c) Note that $\mu = E[X_1] = E[Y]$ and we want $P[|Y - \mu| \geq k] \leq 0.01$.

$$\begin{aligned} P[|Y - \mu| \geq k] &= P[|Y - E[Y]| \geq k] \\ &\leq \frac{Var[Y]}{k^2} \\ &= \frac{Var[X_1]}{nk^2} \\ &\leq \frac{49}{nk^2} \end{aligned}$$

To make this at most 0.01, we need $n \geq 100 \times 49/k^2 = 4 \times 4900$ when $k = 0.5$ year.

d)

$$
\begin{aligned}
P[|Y-\mu| \geq k] &= P\left[\left|\frac{\sum_{m=1}^{n}(X_m-\mu)}{n}\right| \geq k\right] \\
&= P\left[\left|\frac{\sum_{m=1}^{n}(X_m-\mu)}{n\sigma}\right| \geq \frac{k}{\sigma}\right] && [\ \sigma = SD[X_1]\ ] \\
&= P\left[\left|\frac{\sum_{m=1}^{n}(X_m-\mu)}{\sqrt{n}\sigma}\right| \geq \frac{k\sqrt{n}}{\sigma}\right] \\
&\approx P\left[|Z| \geq \frac{k\sqrt{n}}{\sigma}\right] && \text{[by CLT]} \\
&= 2P\left[Z \geq \frac{k\sqrt{n}}{\sigma}\right]
\end{aligned}
$$

where $Z \sim \mathcal{N}(0,1)$. So we want

$$P\left[Z \geq \frac{k\sqrt{n}}{\sigma}\right] \leq 0.005$$

From the table, $P[Z \leq 2.58] \leq 0.9951$, so $P[Z \geq 2.58] \leq 0.005$. So, we want

$$\frac{k\sqrt{n}}{\sigma} \geq 2.58 \quad \Longleftrightarrow \quad n \geq (2.58)^2 \frac{\sigma^2}{k^2}$$

Since we know $Var[X_1] \leq 7^2$, we can take

$$n \geq (2.58)^2 \frac{49}{k^2}$$

So $n = 1305$ is safe value when $k = 0.5$.

Note: If we are willing to do more work in part a), we can show that $Var[X_1] \leq \frac{1}{4} \times 7^2$ year^2. This will reduce the size of n in parts b), c) and d) by a factor of $1/4$.

Here's how to get $Var[X_1] \leq 7^2/4$: Let t be any constant:

$$\begin{aligned}
E[(X_1 - t)^2] &= E[\ ((X_1 - \mu) + (\mu - t))^2\ ] \\
&= E[(X_1 - \mu)^2] + 2E[(X_1 - \mu)(\mu - t)] + E[(\mu - t)^2] \\
&= E[(X_1 - \mu)^2] + 2E[X_1 - \mu]E[\mu - t] + E[(\mu - t)^2] \\
&= E[(X_1 - \mu)^2] + 2 \times 0 \times (\mu - t) + (\mu - t)^2 \\
&= \underbrace{E[(X_1 - \mu)^2]}_{=Var[X_1]} + \underbrace{(\mu - t)^2}_{\geq 0}
\end{aligned}$$

So, we have that for any constant t:

$$Var[X_1] \leq E[(X_1 - t)^2]$$

and choose $t = \frac{7}{2}$:

$$Var[X_1] \leq E[(X_1 - \frac{7}{2})^2]$$

Since $0 \leq X_1 \leq 7$, then $-\frac{7}{2} \leq X_1 - \frac{7}{2} \leq \frac{7}{2}$ and $(X_1 - \frac{7}{2})^2 \leq 7^2/4$:

$Var[X_1] \leq E[(X_1 - 7/2)^2] \leq E[7^2/4] = 7^2/4$