7 Topics: Moment generating functions, conditional distribution and conditional expectation

7.1 Prerequisites: Lecture 18

Exercice 7- 1: Use moment generating functions to find the mean and variance of

(a) $X \sim \mathrm{Poi}(\lambda)$,

(b) $X \sim \mathrm{Bin}(n, p)$,

(c) $X \sim \mathrm{Exp}(\lambda)$,

(d) $X \sim \mathrm{N}(\mu, \sigma^2)$.

Solution:

(a) $X \sim \mathrm{Poi}(\lambda)$: For $t \in \mathbb{R}$, we have

$$M_X(t) = \sum_{n=0}^{\infty} e^{tn} \frac{\lambda^n}{n!} e^{-\lambda} = \sum_{n=0}^{\infty} \frac{(e^t\lambda)^n}{n!} e^{-\lambda} = \exp(-\lambda + \lambda e^t) = \exp(\lambda(e^t - 1)).$$

Then

$$M'_X(t) = \exp(\lambda(e^t - 1))\lambda e^t, \qquad M''_X(t) = \exp(\lambda(e^t - 1))\lambda e^t + \exp(\lambda(e^t - 1))(\lambda e^t)^2,$$

and hence $\mathrm{E}(X) = M'_X(0) = \lambda$, $\mathrm{E}(X^2) = M''_X(0) = \lambda + \lambda^2$ and $\mathrm{Var}(X) = \lambda + \lambda^2 - \lambda^2 = \lambda$.

(b) $X \sim \mathrm{Bin}(n, p)$: For $t \in \mathbb{R}$, we have

$$M_X(t) = \sum_{j=0}^{n} e^{jt} \binom{n}{j} p^j (1-p)^{n-j} = \sum_{j=0}^{n} \binom{n}{j} (e^t p)^j (1-p)^{n-j} = (e^t p + 1 - p)^n.$$

Then

$$M'_X(t) = n(e^t p + 1 - p)^{n-1} p e^t,$$
$$M''_X(t) = n(n-1)(e^t p + 1 - p)^{n-2}(pe^t)^2 + n(e^t p + 1 - p)^{n-1} p e^t,$$

hence $\mathrm{E}(X) = M'_X(0) = np$ and $\mathrm{E}(X^2) = M''_X(0) = np + n(n-1)p^2$, hence $\mathrm{Var}(X) = np(1-p)$.

(c) $X \sim \mathrm{Exp}(\lambda)$: For $t < \lambda$, we have

$$M_X(t) = \int_0^{\infty} e^{tx} \lambda e^{-\lambda x} dx = \frac{\lambda}{\lambda - t}.$$

Then

$$M'_X(t) = \frac{\lambda}{(\lambda - t)^2}, \qquad M''_X(t) = \frac{2\lambda}{(\lambda - t)^3},$$

hence $\mathrm{E}(X) = M'_X(0) = 1/\lambda$ and $\mathrm{E}(X^2) = M''_X(0) = 2/\lambda^2$, hence $\mathrm{Var}(X) = 1/\lambda^2$.

(d) $X \sim \mathrm{N}(\mu, \sigma^2)$. From the lectures we know that, for $t \in \mathbb{R}$, we have

$$M_X(t) = \int_{-\infty}^{\infty} e^{tx} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} dx = e^{\mu t + \sigma^2 t^2/2}.$$

Then

$$M'_X(t) = e^{\mu t + \sigma^2 t^2/2}(\mu + \sigma^2 t), \qquad M''_X(t) = e^{\mu t + \sigma^2 t^2/2}(\mu + \sigma^2 t)^2 + e^{\mu t + \sigma^2 t^2/2}\sigma^2,$$

hence $\mathrm{E}(X) = M'_X(0) = \mu$ and $\mathrm{E}(X^2) = M''_X(0) = \mu^2 + \sigma^2$, hence $\mathrm{Var}(X) = \sigma^2$.

Exercice 7- 2: (Suggested for personal/peer tutorial) Use moment generating functions to prove that for independent random variables $X \sim \mathrm{N}(\mu_X, \sigma_X^2), Y \sim \mathrm{N}(\mu_Y, \sigma_Y^2)$, we have that $X + Y \sim \mathrm{N}(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2)$.

(a)

Solution: From Theorem 12.2.7 we deduce that, for $t \in \mathbb{R}$,

$$M_{X+Y}(t) = M_X(t)M_Y(t) = e^{\mu_X t + \sigma_X^2 t^2/2} e^{\mu_Y t + \sigma_Y^2 t^2/2} = e^{(\mu_X + \mu_Y)t + \frac{1}{2}(\sigma_X^2 + \sigma_Y^2)t^2}.$$

We observe that the right hand side is the m.g.f. of a random variable with $\mathrm{N}(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2)$ distribution. Then Theorem 12.2.8 allows us to conclude.

Exercice 7- 3: Suppose that X_1 and X_2 are independent and identically distributed random variables, each having a standard normal distribution. Let random variable V be defined by

$$V = X_1^2 + X_2^2.$$

Find the pdf of V.

Solution: There are many possible routes for solving this problem.

(a) The possibly simplest approach is to use m.g.f.s. We have from Exercise 5- 8 that if $X \sim \mathrm{N}(0, 1)$ then

$$Y = X^2 \sim \mathrm{Gamma}\left(\frac{1}{2}, \frac{1}{2}\right)$$

and hence

$$\begin{aligned} M_Y(t) &= \int_0^\infty e^{ty} \left(\frac{1}{2\pi}\right)^{1/2} y^{-1/2}\, e^{-y/2}\, dy = \left(\frac{1}{2\pi}\right)^{1/2} \int_0^\infty y^{-1/2}\, e^{-y(1-2t)/2}\, dy \\ &= \left(\frac{1}{2\pi}\right)^{1/2} \frac{(2\pi)^{1/2}}{(1-2t)^{1/2}} = \frac{1}{(1-2t)^{1/2}}, \end{aligned}$$

for $\frac{1}{2} > t$, as the integrand is proportional to a $\mathrm{Gamma}(1/2, (1-2t)/2)$ pdf. Hence

$$M_Y(t) = \left(\frac{1}{1-2t}\right)^{1/2}.$$

Now, let $Y_1 = X_1^2$ and $Y_2 = X_2^2$, so that Y_1 and Y_2 are independent $\mathrm{Gamma}(1/2, 1/2)$ variables. Now we have from a key m.g.f. result, for $\frac{1}{2} > t$,

$$V = Y_1 + Y_2 \quad \Longrightarrow \quad M_V(t) = M_{Y_1}(t)M_{Y_2}(t) = \left(\frac{1}{1-2t}\right)^{1/2}\left(\frac{1}{1-2t}\right)^{1/2} = \left(\frac{1}{1-2t}\right),$$

and hence, noting that this is the mgf of a Gamma random variable with parameters 1 and $\frac{1}{2}$, we conclude that

$$V \sim \mathrm{Gamma}\left(1, \frac{1}{2}\right) = \mathrm{Gamma}\left(\frac{2}{2}, \frac{1}{2}\right) \equiv \chi_2^2.$$

(b) By using a joint p.d.f. approach, we could also write

$$F_V(v) = \mathrm{P}\left(V \leq v\right) = \mathrm{P}\left(X_1^2 + X_2^2 \leq v\right) = \int_{A_v}\int f_{X_1}(x_1) f_{X_2}(x_2)\, dx_1 dx_2,$$

where

$$A_v = \left\{(x_1, x_2) \;:\; x_1^2 + x_2^2 \leq v\right\}$$

that is, the integral over the region A_v of the joint density function of X_1 and X_2. Now, should reparameterize into polar coordinates in the double integral; let $x_1 = r\cos\theta$ and $x_2 = r\sin\theta$. Then

$$\begin{aligned}
F_V(v) &= \int_{A_v}\int f_{X_1}(x_1) f_{X_2}(x_2)\, dx_1 dx_2 = \int_{A_v}\int \frac{1}{2\pi} e^{-(x_1^2+x_2^2)/2}\, dx_1 dx_2 \\
&= \int_{r=0}^{\sqrt{v}} \int_0^{2\pi} \frac{1}{2\pi} e^{-r^2/2} r\, dr d\theta \\
&= \int_{r=0}^{\sqrt{v}} r e^{-r^2/2}\, dr = 1 - e^{-v/2}, \quad v > 0
\end{aligned}$$

so $V \sim \mathrm{Exp}(1/2) \equiv \chi_2^2$.

(c) We could also use the convolution theorem: Let $V = Y_1 + Y_2$ where $Y_1 = X_1^2$ and $Y_2 = X_2^2$. We have already shown that $Y_1 \sim \mathrm{Gamma}(1/2, 1/2)$ so $f_{Y_i}(y) = (2\pi)^{-1/2} y^{-1/2} e^{-y/2},\ y \geq 0$. The range of V is $(0, \infty)$. The convolution theorem is:

$$\begin{aligned}
f_V(v) &= \int_{-\infty}^{\infty} f_{Y_1,Y_2}(y, v-y)\, dy \\
&= \int_0^v f_{Y_1}(y) f_{Y_2}(v-y)\, dy \ \text{ from independence} \\
&= \int_0^v \frac{1}{\sqrt{2\pi}} y^{-1/2} e^{-y/2} \frac{1}{\sqrt{2\pi}} (v-y)^{-1/2} e^{-(v-y)/2}\, dy \\
&= \frac{1}{2\pi} e^{-v/2} \int_0^v y^{-1/2}(v-y)^{-1/2}\, dy,\ v > 0.
\end{aligned}$$

Note that the kernel of this integral looks like a scaled Beta distribution. Substitute $z = y/v$ (could also use a trig substitution) then

$$\begin{aligned}
f_V(v) &= \frac{1}{2\pi} e^{-v/2} \int_0^1 (vz)^{-1/2}(v - vz)^{-1/2}\, v dz \\
&= \frac{1}{2\pi} e^{-v/2} \int_0^1 (z)^{-1/2}(1-z)^{-1/2}\, \mathrm{d}z \\
&= \frac{1}{2\pi} e^{-v/2} \frac{\Gamma(1/2)\Gamma(1/2)}{\Gamma(1)} = \frac{1}{2\pi} e^{-v/2} \sqrt{\pi}\sqrt{\pi} \\
&= \frac{1}{2} e^{-v/2},\ v > 0,
\end{aligned}$$

which we recognise as the p.d.f. of an $\mathrm{Exp}(1/2)$ random variable.

7.2 Prerequisites: Lecture 19

Exercice 7- 4: Consider tossing a coin repeatedly, where the probability of heads appearing in one toss is given by $p \in (0, 1)$. Let X denote the length of the initial run (i.e. if you toss heads first, how many heads do you toss before tossing tail and vice versa if you toss tail first). By conditioning on the outcome of the first coin toss and by using the law of total expectation, find $\mathrm{E}(X)$.

Solution: Let H denote the event that the first coin toss gives heads. Then H^c is the event that the first coin toss gives tail. Then, for $x \in \mathbb{N}$, we have

$$\mathrm{P}(X = x|H) = p^{x-1}(1-p),$$

since after the initial toss of heads we need to toss heads $x-1$ times and then toss tail ones to obtain a run of heads of length x.

Similarly, we get for $x \in \mathbb{N}$,

$$\mathrm{P}(X = x|H^c) = (1-p)^{x-1}p.$$

By the law of total (noting that H and H^c form a partition of the sample space), we get

$$\mathrm{E}(X) = p\mathrm{E}(X|H) + (1-p)\mathrm{E}(X|H^c).$$

Since

$$\mathrm{E}(X|H) = \sum_x x\mathrm{P}(X = x|H) = \sum_x xp^{x-1}(1-p) = (1-p)\sum_x p^{x-1}.$$

Now we can use a little trick: Note that since $p \in (0, 1)$, we know that the geometric series is given by

$$\sum_{x=0}^{\infty} p^x = \frac{1}{1-p} = f(p).$$

So, if we view this as a function in p and differentiate both the left and the right hand side (assuming we can interchange the infinite sum and the differentiation), then

$$f'(p) = \sum_{x=1}^{\infty} xp^{x-1} = \frac{1}{(1-p)^2}.$$

Hence

$$\mathrm{E}(X|H) = \sum_x x\mathrm{P}(X = x|H) = \sum_x xp^{x-1}(1-p) = \frac{(1-p)}{(1-p)^2} = \frac{1}{1-p}.$$

Using the same arguments, we get

$$\mathrm{E}(X|H^c) = \sum_x x\mathrm{P}(X = x|H) = \sum_x x(1-p)^{x-1}p = p\frac{1}{(1-(1-p))^2} = \frac{1}{p}.$$

So, altogether we have

$$\begin{aligned}\mathrm{E}(X) &= p\mathrm{E}(X|H) + (1-p)\mathrm{E}(X|H^c). = p\frac{1}{1-p} + (1-p)\frac{1}{p} = \frac{p^2 + (1-p)^2}{(1-p)p} \\ &= \frac{p^2 + 1 - 2p + p^2}{(1-p)p} = \frac{2p^2 + 1 - 2p}{(1-p)p} = -2 + \frac{1}{p(1-p)}.\end{aligned}$$