170E Week 9 Discussion Notes

Colin Ni

November 24, 2025

Exercise. (Book 4.3.11) Suppose $X \sim \mathrm{Geom}(p)$, and suppose Y is a random variable such that the conditional distribution $Y \mid X = x \sim \mathrm{Pois}(x)$. Find $\mathbb{E}(Y)$ and $\mathrm{Var}(Y)$.

Solution. We compute

$$\mathbb{E}(Y) = \mathbb{E}(\mathbb{E}(Y \mid X)) = \mathbb{E}(\mathrm{Pois}(X)) = \mathbb{E}(X) = \frac{1}{p}$$

by conditioning on X. Note that, roughly speaking, the randomness in the random variable $\mathbb{E}(Y \mid X)$ is coming from X, so the outer $\mathbb{E}$ is being taken over X.

To compute $\mathrm{Var}(Y)$, we will use the so-called *law of total variance*:

$$\mathrm{Var}(Y) = \mathbb{E}(\mathrm{Var}(Y \mid X)) + \mathrm{Var}(\mathbb{E}(Y \mid X)).$$

This is easy to derive:

$$\begin{aligned}
\mathrm{Var}(Y) &= \mathbb{E}(Y^2) - \mathbb{E}(Y)^2 \\
&= \mathbb{E}(\mathbb{E}(Y^2 \mid X)) - \mathbb{E}(\mathbb{E}(Y \mid X))^2 && \text{(condition on } X \text{ twice)} \\
&= \mathbb{E}(\mathrm{Var}(Y \mid X)) + \mathbb{E}(\mathbb{E}(Y \mid X)^2) - \mathbb{E}(\mathbb{E}(Y \mid X))^2 && \text{(def of variance)} \\
&= \mathbb{E}(\mathrm{Var}(Y \mid X)) + \mathrm{Var}(\mathbb{E}(Y \mid X)). && \text{(def of variance)}
\end{aligned}$$

Here, we have that

$$\mathbb{E}(\mathrm{Var}(Y \mid X)) = \mathbb{E}(\mathrm{Var}(\mathrm{Pois}(X))) = \mathbb{E}(X) = \frac{1}{p}$$

and

$$\mathrm{Var}(\mathbb{E}(Y \mid X)) = \mathrm{Var}(\mathbb{E}(\mathrm{Pois}(X))) = \mathrm{Var}(X) = \frac{1-p}{p^2},$$

so

$$\mathrm{Var}(Y) = \frac{1}{p} + \frac{1-p}{p^2} = \frac{1}{p^2}.$$ □

Exercise. (Book 4.4.20) Let $X \sim \mathrm{Unif}(1, 2)$, and let Y be a random variable such that the conditional pdf of $Y \mid X = x$ is

$$f_{Y \mid X=x}(y) = (x+1)y^x \quad \text{for} \quad 0 < y < 1.$$

Find $\mathbb{E}(Y)$.

Solution. Since

$$\begin{aligned}\mathbb{E}(Y \mid X = x) &= \int_0^1 y f_{Y|X=x}(y)\,\mathrm{d}y \\ &= \int_0^1 (x+1)y^{x+1}\,\mathrm{d}y \\ &= \frac{x+1}{x+2}y^{x+2}\Big|_0^1 \\ &= \frac{x+1}{x+2},\end{aligned}$$

we have

$$\mathbb{E}(Y \mid X) = \frac{X+1}{X+2}.$$

Thus

$$\begin{aligned}\mathbb{E}(Y) &= \mathbb{E}(\mathbb{E}(Y \mid X)) \\ &= \mathbb{E}\left(\frac{X+1}{X+2}\right) \\ &= \int_1^2 \frac{x+1}{x+2} f_X(x)\,\mathrm{d}x \\ &= \int_1^2 \frac{(x+2)-1}{x+2}\,\mathrm{d}x \\ &= \int_1^2 \left(1 - \frac{1}{x+2}\right)\,\mathrm{d}x \\ &= 1 - \ln(x+2)\Big|_1^2\,\mathrm{d}x \\ &= 1 - \ln(4) + \ln(3).\end{aligned}$$

□

Last week during discussion section we repeatedly used the following fact:

Exercise. Let $X \sim \mathrm{Norm}(\mu, \Sigma)$ be a multivariate normal random variable in n variables, and let A be an invertible $n \times n$ matrix and $b \in \mathbb{R}^n$ be a vector. Then

$$AX + b \sim \mathrm{Norm}(A\mu + b, A\Sigma A^T).$$

Solution. Recall that the pdf of X is given by

$$f_X(x) = \frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{|\Sigma|}} \exp\left(-\frac{1}{2}(x-\mu)^T\Sigma^{-1}(x-\mu)\right).$$

We will use the change of variables theorem to determine the pdf of $Y = AX+b$. Our inverse function is $X = A^{-1}(Y-b)$, which has Jacobian A^{-1}. Thus

$$\begin{aligned}f_Y(y) &= |A^{-1}|f_X(A^{-1}(y-b)) \\ &= |A^{-1}|\frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{|\Sigma|}} \exp\left(-\frac{1}{2}(A^{-1}(y-b)-\mu)^T\Sigma^{-1}(A^{-1}(y-b)-\mu)\right).\end{aligned}$$

The constant is

$$\frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{|A\Sigma A^T|}}$$

because

$$\frac{|A^{-1}|}{\sqrt{|\Sigma|}} = \frac{1}{|A|\sqrt{|\Sigma|}} = \frac{1}{\sqrt{|A\Sigma A^T|}}.$$

For the expression in the exp, observe

$$\begin{aligned}
&-\frac{1}{2}(A^{-1}(y-b)-\mu)^T\Sigma^{-1}(A^{-1}(y-b)-\mu) \\
&\quad = -\frac{1}{2}(A^{-1}(y-(A\mu+b)))^T\Sigma^{-1}A^{-1}(y-(A\mu+b)) \\
&\quad = -\frac{1}{2}(y-(A\mu+b))^T(A^{-1})^T\Sigma^{-1}A^{-1}(y-(A\mu+b)) \\
&\quad = -\frac{1}{2}(y-(A\mu+b))^T(A\Sigma A^T)^{-1}(y-(A\mu+b)).
\end{aligned}$$

Thus the pdf is

$$\frac{1}{(2\pi)^{\frac{n}{2}}\sqrt{|A\Sigma A^T|}}\exp\left(-\frac{1}{2}(y-(A\mu+b))^T(A\Sigma A^T)^{-1}(y-(A\mu+b))\right)$$

which is the pdf of

$$\mathrm{Norm}(A\mu+b, A\Sigma A^T),$$

as required. $\square$

We used this fact to show that given a bivariate normal

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim \mathrm{Norm}(\mu, \Sigma),$$

we could make the following transformations:

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim \mathrm{Norm}(\mu, \Sigma)$$

$$\Big\downarrow \text{standardize}$$

$$\begin{pmatrix} Z_X \\ Z_Y \end{pmatrix} = \begin{pmatrix} \frac{1}{\sigma_X} & 0 \\ 0 & \frac{1}{\sigma_Y} \end{pmatrix}\begin{pmatrix} X \\ Y \end{pmatrix} - \begin{pmatrix} \frac{\mu_X}{\sigma_X} \\ \frac{\mu_Y}{\sigma_Y} \end{pmatrix} \sim \mathrm{Norm}\left(0, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right)$$

$$\Big\downarrow \text{remove correlation}$$

$$\begin{pmatrix} U_X \\ U_Y \end{pmatrix} = \frac{1}{\sqrt{1-\rho^2}}\begin{pmatrix} 1 & -\rho \\ 0 & \sqrt{1-\rho^2} \end{pmatrix}\begin{pmatrix} Z_X \\ Z_Y \end{pmatrix} \sim \mathrm{Norm}\left(0, \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right).$$

Exercise. (Lecture notes 4.3) Given a bivariate normal

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim \mathrm{Norm}(\mu, \Sigma),$$

find the conditional distribution $X \mid Y = y$ and the conditional expectation $\mathbb{E}(X \mid Y = y)$.

Solution. Writing our bivariate normal in terms of U_X, U_Y gives

$$\begin{aligned}\begin{pmatrix} X \\ Y \end{pmatrix} &= \begin{pmatrix} \sigma_X & 0 \\ 0 & \sigma_Y \end{pmatrix} \begin{pmatrix} Z_X \\ Z_Y \end{pmatrix} + \begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix} \\ &= \begin{pmatrix} \sigma_X & 0 \\ 0 & \sigma_Y \end{pmatrix} \begin{pmatrix} \sqrt{1-\rho^2} & \rho \\ 0 & 1 \end{pmatrix} \begin{pmatrix} U_X \\ U_Y \end{pmatrix} + \begin{pmatrix} \mu_X \\ \mu_Y \end{pmatrix} \\ &\begin{pmatrix} \sigma_X \left(\sqrt{1-\rho^2} U_X + \rho U_Y \right) + \mu_X \\ \sigma_Y U_Y + \mu_Y \end{pmatrix}.\end{aligned}$$

If $Y = y$, then

$$U_Y = \frac{y - \mu_Y}{\sigma_Y},$$

so

$$\begin{aligned}(X \mid Y = y) &= \sigma_X \left(\sqrt{1-\rho^2} U_X + \rho U_Y \right) + \mu_X \\ &= \sigma_X \left(\sqrt{1-\rho^2} U_X + \rho \frac{y - \mu_Y}{\sigma_Y} \right) + \mu_X \\ &\sim \operatorname{Norm} \left(\frac{\rho \sigma_X}{\sigma_Y} (y - \mu_Y) + \mu_X, (1-\rho^2) \sigma_X^2 \right).\end{aligned}$$

We can of course easily read off the conditional expectation. □

Remark. More generally, one can determine the conditional distribution of a multivariate normal given the values of some of its variables: see this (download and click).