Ryan Ta

Theorem 2.3.1

Note: This is Theorem 2.3.1 (page 98) of the textbook *Introduction to Mathematical Statistics* (seventh edition) by Robert V. Hogg, Joseph W. McKean, Allen T. Craig. I am following the proof of Theorem 2.3.1 but filling in intermediate steps here, so that the proof is hopefully easier to read.

Theorem (Theorem 2.3.1 of Hogg, McKean, Craig). *Let (X_1, X_2) be a random vector such that the variance of X_2 is finite. Then we have:*

(a) $E[E(X_2|X_1)] = E(X_2)$,

(b) $\mathrm{Var}[E(X_2|X_1)] \leq \mathrm{Var}(X_2)$.

Proof. First, we will prove statement (a). For the continuous case, we have

$$\begin{aligned} E(X_2) &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_2 f_{X_1,X_2}(x_1,x_2)\,dx_2\,dx_1 \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_2 f_{X_1,X_2}(x_1,x_2)\frac{f_{X_1}(x_1)}{f_{X_1}(x_1)}\,dx_2\,dx_1 \\ &= \int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty} x_2\frac{f_{X_1,X_2}(x_1,x_2)}{f_{X_1}(x_1)}\,dx_2\right) f_{X_1}(x_1)\,dx_1 \\ &= \int_{-\infty}^{\infty} E(X_2|x_1) f_{X_1}(x_1)\,dx_1 \\ &= E[E(X_2|X_1)]. \end{aligned}$$

For the discrete case, we have

$$\begin{aligned} E(X_2) &= \sum_{x_1}\sum_{x_2} x_2 p_{X_1,X_2}(x_1,x_2) \\ &= \sum_{x_1}\sum_{x_2} x_2 p_{X_1,X_2}(x_1,x_2)\frac{p_{X_1}(x_1)}{p_{X_1}(x_1)} \\ &= \sum_{x_1}\left(\sum_{x_2} x_2\frac{p_{X_1,X_2}(x_1,x_2)}{p_{X_1}(x_1)}\right) p_{X_1}(x_1) \\ &= \sum_{x_1} E(X_2|x_1) p_{X_1}(x_1) \\ &= E[E(X_2|X_1)]. \end{aligned}$$

This completes our proof of statement (a).

Next, we will prove statement (b). We recall that expectation of a variable is the mean; for instance, we consider $\mu_2 = E(X_2)$. Then, by the linearity of expectation, we have

$$\begin{aligned} \mathrm{Var}(X_2) &= E[(X_2-\mu_2)^2] \\ &= E[((X_2 - E(X_2|X_1)) + (E(X_2|X_1)-\mu_2))^2] \\ &= E[(X_2 - E(X_2|X_1))^2 + 2(X_2 - E(X_2|X_1))(E(X_2|X_1)-\mu_2) + (E(X_2|X_1)-\mu_2)^2] \\ &= E[(X_2 - E(X_2|X_1))^2] + 2E[(X_2 - E(X_2|X_1))(E(X_2|X_1)-\mu_2)] + E[(E(X_2|X_1)-\mu_2)^2]. \end{aligned}$$

We also note that, since we notice $(X_2 - E(X_2|X_1))^2 \geq 0$ and $f_{X_1,X_2}(x_1,x_2) \geq 0$, it follows that the first term $E[(X_2 - E(X_2|X_1)^2]$ in our latest expression of $\mathrm{Var}(X_2)$ satisfies

$$\begin{aligned} E[(X_2 - E(X_2|X_1)^2] &= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x_2 - E(X_2|x_1))^2 f_{X_1,X_2}(x_1,x_2)\,dx_2\,dx_1 \\ &\geq \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} 0\,dx_2\,dx_1 \\ &= 0. \end{aligned}$$

Therefore, we have

$$\begin{aligned} \mathrm{Var}(X_2) &= E[(X_2 - E(X_2|X_1))^2] + 2E[(X_2 - E(X_2|X_1))(E(X_2|X_1)-\mu_2)] + E[(E(X_2|X_1)-\mu_2)^2] \\ &= E[(X_2 - E(X_2|X_1))^2] + 2(0) + E[(E(X_2|X_1)-\mu_2)^2] \\ &\geq 0 + 2(0) + E[(E(X_2|X_1)-\mu_2)^2] \\ &= E[(E(X_2|X_1)-\mu_2)^2] \\ &= \mathrm{Var}(E(X_2|X_1)), \end{aligned}$$

as desired. □