1. Question: Conditional expectation (*elementary*)

1.1. Suppose we draw $X \sim \text{Unif}(0,1)$. After we observe $X = x$, we draw $Y \mid X = x \sim \text{Unif}(x, 1)$, resulting in the conditional density $f_{Y|X}(y \mid x) = 1/(1-x)$ for $x < y < 1$.
Find the conditional expectation of Y given X.

1.2. Let $X \sim$ Uniform $(0,1)$. Let $0 < a < b < 1$. Consider

$$Y = \begin{cases} 1 & 0 < x < b \\ 0 & \text{otherwise,} \end{cases} \quad \text{and} \quad Z = \begin{cases} 1 & a < x < 1 \\ 0 & \text{otherwise.} \end{cases}$$

(i) Are Y and Z independent? Why/Why not?

(ii) Find $\mathbb{E}(Y|Z)$. Hint: What values z can Z take? Now find $\mathbb{E}(Y \mid Z = z)$.

1.3. Let $r(x)$ be a function of x and let $s(y)$ be a function of y. Show that

$$\mathbb{E}[r(X)s(Y) \mid X] = r(X)\mathbb{E}[s(Y) \mid X].$$

Solution:

1.1.

$$\mathbb{E}(Y \mid X = x) = \int_x^1 y f_{Y|X}(y \mid x) dy = \frac{1}{1-x} \int_x^1 y dy = \frac{1+x}{2},$$

Thus, $\mathbb{E}(Y \mid X) = (1+X)/2$. Notice that $\mathbb{E}(Y \mid X) = (1+X)/2$ is a random variable whose value is the number $\mathbb{E}(Y \mid X = x) = (1+x)/2$ once $X = x$ is observed.

1.2. (i) X and Z are not independent, because

$$\begin{aligned} &\mathbb{P}(Y = 1) = \mathbb{P}(x < b) = b \\ &\mathbb{P}(Z = 1) = \mathbb{P}(x > a) = 1 - a \\ &\mathbb{P}(Y = 1, Z = 1) = \mathbb{P}(a < x < b) = b - a \\ &\mathbb{P}(Y = 1)\mathbb{P}(Z = 1) \neq \mathbb{P}(Y = 1, Z = 1). \end{aligned}$$

(ii) If $Z = 0$: $x < a \Longrightarrow x < b \Longrightarrow Y = 1 \Longrightarrow \mathbb{E}[Y|Z] = \mathbb{E}[Y] = 1$. Meanwhile,
if $Z = 1$: $x > a \Longrightarrow \mathbb{P}(Y = 1) = \mathbb{P}(a < x < b \mid a < x < 1) = \frac{b-a}{1-a} \Longrightarrow \mathbb{E}(Y \mid Z = 1) = \dfrac{b-a}{1-a}$.

$$\Rightarrow \mathbb{E}[Y \mid Z] = \begin{cases} 1 & Z = 0 \\ \frac{b-a}{1-a} & Z = 1 \end{cases}$$

1.3.

$$\begin{aligned} \mathbb{E}[r(X)s(Y) \mid X] = \int r(X)s(y) d\mathbb{P}_{Y|X}(y) &= \int r(X)s(y)f(y \mid x)dy \\ &= r(X) \int s(y)f(y \mid x)dy \\ &= r(X)\mathbb{E}[s(Y) \mid X]. \end{aligned}$$

2. Question: Conditional variance and law of total variance (*medium*)

2.1. The Bernoulli distribution with parameter p is defined via the pmf $f(x) = p^x(1-p)^{1-x}, \quad x \in \{0, 1\}$.
Consider two random variables $X, Y \sim \text{Bernoulli}\left(\frac{2}{5}\right)$ with

$$X \mid Y = 0 \sim \text{ Bernoulli}\left(\frac{2}{3}\right), \quad P(X = 0 \mid Y = 1) = 1, \quad \text{Var}(\mathbb{E}[X|Y]) = \frac{8}{75}.$$

Find the pmf of $V := \mathrm{Var}(X|Y)$, $\mathbb{E}[V]$, and check that $\mathrm{Var}(X) = \mathbb{E}[V] + \mathrm{Var}(\mathbb{E}[X|Y])$.

2.2. Consider a random variable N that takes values in $\mathbb{N}$ and suppose that we know $\mathbb{E}[N]$ and $\mathrm{Var}(N)$. Find the expectation and variance of the random variable

$$Y = \sum_{i=1}^{N} X_i,$$

where the X_i are i.i.d. and also independent of N.
Hint: You may use the fact that for independent X, Y, Z, we have $\mathbb{E}[X + Y|Z] = \mathbb{E}[X|Z] + \mathbb{E}[Y|Z] = \mathbb{E}[X] + \mathbb{E}[Y]$, with both equality following immediately from independence.

Solution:

2.1.

- To find the pmf of V, we note that V is a function of Y. Specifically,

$$V = \mathrm{Var}(X \mid Y) = \begin{cases} \mathrm{Var}(X \mid Y=0) & \text{if } Y=0 \\ \mathrm{Var}(X \mid Y=1) & \text{if } Y=1 \end{cases}$$

Therefore,

$$V = \mathrm{Var}(X \mid Y) = \begin{cases} \mathrm{Var}(X \mid Y=0) & \text{with probability } \frac{3}{5} \\ \mathrm{Var}(X \mid Y=1) & \text{with probability } \frac{2}{5} \end{cases}$$

Now, since $X \,\big|\, Y = 0 \sim \text{Bernoulli}\left(\frac{2}{3}\right)$, we have

$$\mathrm{Var}(X \mid Y=0) = \frac{2}{3} \cdot \frac{1}{3} = \frac{2}{9}$$

and since given $Y = 1, X = 0$, we have

$$\mathrm{Var}(X \mid Y=1) = 0$$

Thus,

$$V = \mathrm{Var}(X \mid Y) = \begin{cases} \frac{2}{9} & \text{with probability } \frac{3}{5} \\ 0 & \text{with probability } \frac{2}{5}. \end{cases}$$

So, we can write

$$P_V(v) = \begin{cases} \frac{3}{5} & \text{if } v = \frac{2}{9} \\ \frac{2}{5} & \text{if } v = 0 \\ 0 & \text{otherwise.} \end{cases}$$

- To find $\mathbb{E}[V]$, we write

$$\mathbb{E}[V] = \frac{2}{9} \cdot \frac{3}{5} + 0 \cdot \frac{2}{5} = \frac{2}{15}.$$

- To check that $\mathrm{Var}(X) = \mathbb{E}[V] + \mathrm{Var}(\mathbb{E}[X|Y])$, we just note that

$$\mathrm{Var}(X) = \frac{2}{5} \cdot \frac{3}{5} = \frac{6}{25}, \quad \mathbb{E}[V] = \frac{2}{15}, \quad \mathrm{Var}(\mathbb{E}[X|Y]) = \frac{8}{75},$$

and $\frac{2}{15} + \frac{8}{75} = \frac{18}{75} = \frac{6}{25}$.

2.2. To find $\mathbb{E}[Y]$, we cannot directly use the linearity of expectation because N is random. But, conditioned on $N = n$, we can use linearity and find $\mathbb{E}[Y \mid N = n]$; so, we use the rule of iterated expectations:

$$\mathbb{E}[Y] = \mathbb{E}[\mathbb{E}[Y \mid N]] = E\left[E\left[\sum_{i=1}^{N} X_i \mid N\right]\right] = E\left[\sum_{i=1}^{N} E\left[X_i \mid N\right]\right]$$

$$= E\left[\sum_{i=1}^{N} E\left[X_i\right]\right] = \mathbb{E}[N\mathbb{E}[X]] = \mathbb{E}[X]\mathbb{E}[N].$$

To find $\operatorname{Var}(Y)$, we use the law of total variance:

$$\begin{aligned}\operatorname{Var}(Y) &= \mathbb{E}[\operatorname{Var}(Y \mid N)) + \operatorname{Var}(\mathbb{E}[Y \mid N]) \\ &= \mathbb{E}[\operatorname{Var}(Y \mid N)) + \operatorname{Var}(N\mathbb{E}[X]) \qquad (\star) \\ &= \mathbb{E}[\operatorname{Var}(Y \mid N)) + (\mathbb{E}[X])^2 \operatorname{Var}(N).\end{aligned}$$

To find $\mathbb{E}[\operatorname{Var}(Y \mid N))$, note that, given $N = n, Y$ is a sum of n independent random variables. Thus, we can write

$$\begin{aligned}\operatorname{Var}(Y \mid N) &= \sum_{i=1}^{N} \operatorname{Var}(X_i \mid N) \\ &= \sum_{i=1}^{N} \operatorname{Var}(X_i) \qquad (\text{since } X_i \text{ 's are independent of } N\ ) \\ &= N \operatorname{Var}(X).\end{aligned}$$

Therefore, we have

$$\mathbb{E}[\operatorname{Var}(Y \mid N)) = \mathbb{E}[N] \operatorname{Var}(X). \qquad (\star\star)$$

Combining Equations $(\star)$ and $(\star\star)$, we obtain

$$\operatorname{Var}(Y) = \mathbb{E}[N] \operatorname{Var}(X) + (\mathbb{E}[X])^2 \operatorname{Var}(N).$$

3. Question: Properties of the conditional expectation (*slightly advanced*)

Let $X, Y : (\Omega, \mathcal{F}) \longrightarrow (\Omega', \mathcal{F}')$ be random variables with $\mathbb{E}[X], \mathbb{E}[Y] < \infty$.

3.1. Show that if a and b are constants and $\mathcal{A} \subset \mathcal{F}$, then $E(aX + bY \mid \mathcal{A}) = aE(X \mid \mathcal{A}) + bE(X \mid \mathcal{A})$ a.s.

3.2. Show that if $X \leq Y$ a.s., then, for $\mathcal{A} \subset \mathcal{F}$, $E(X \mid \mathcal{A}) \leq E(Y \mid \mathcal{A})$ a.s.
Hint: This can be accomplished by showing that $P(\{E(X \mid \mathcal{A}) > E(Y \mid \mathcal{A})\}) = 0$.

3.3. Let $\mathcal{A}$ and $\mathcal{A}_0$ be σ-algebras satisfying $\mathcal{A}_0 \subset \mathcal{A} \subset \mathcal{F}$. Show that

$$E\left[E(X \mid \mathcal{A}) \mid \mathcal{A}_0\right] = E\left(X \mid \mathcal{A}_0\right) = E\left[E\left(X \mid \mathcal{A}_0\right) \mid \mathcal{A}\right] \text{ a.s.}$$

Solution:

3.1. Note that $aE(X|\mathcal{A}) + bE(Y|\mathcal{A})$ is measurable from $(\Omega, \mathcal{A})$ to $(\Omega, \mathcal{F}')$. For any $A \in \mathcal{A}$, by the linearity of integration,

$$\begin{aligned}\int_A (aX + bY) dP &= a \int_A X dP + b \int_A Y dP \\ &= a \int_A E(X|\mathcal{A}) dP + b \int_A E(Y|\mathcal{A}) dP \\ &= \int_A [aE(X|\mathcal{A}) + bE(Y|\mathcal{A})] dP\end{aligned}$$

By the a.s.-uniqueness of the conditional expectation, $E(aX + bY|\mathcal{A}) = aE(X|\mathcal{A}) + bE(X|\mathcal{A})$ a.s.

3.2. Suppose that $X \leq Y$ a.s. By the definition of the conditional expectation and the property of integration,

$$\int_A E(X \mid \mathcal{A})dP = \int_A X dP \leq \int_A Y dP = \int_A E(Y \mid \mathcal{A})dP$$

where

$$A = \{E(X \mid \mathcal{A}) > E(Y \mid \mathcal{A})\} \in \mathcal{A}.$$

Hence $P(A) = 0$, i.e., $E(X \mid \mathcal{A}) \leq E(Y \mid \mathcal{A})$ a.s.

3.3. Note that $E(X \mid \mathcal{A}_0)$ is measurable from $(\Omega, \mathcal{A}_0)$ to $(\Omega, \mathcal{F}')$ and $\mathcal{A}_0 \subset \mathcal{A}$. Hence $E(X \mid \mathcal{A}_0)$ is measurable from $(\Omega, \mathcal{A})$ to $(\Omega, \mathcal{F}')$ and, thus, $E(X \mid \mathcal{A}_0) = E[E(X \mid \mathcal{A}_0) \mid \mathcal{A}]$ a.s. Since $E[E(X \mid \mathcal{A}) \mid \mathcal{A}_0]$ is measurable from $(\Omega, \mathcal{A}_0)$ to $(\Omega, \mathcal{F}')$ and for any $A \in \mathcal{A}_0 \subset \mathcal{A}$,

$$\int_A E[E(X \mid \mathcal{A}) \mid \mathcal{A}_0] dP = \int_A E(X \mid \mathcal{A})dP = \int_A X dP$$

we conclude that $E[E(X \mid \mathcal{A}) \mid \mathcal{A}_0] = E(X \mid \mathcal{A}_0)$ a.s.

If you have any questions or feedback, please feel free to contact me via E-mail at hannah.kuempel@stat.uni-muenchen.de!!

Also, thank you to H. Pishro-Nik, Jun Shao, and the authors of the book *All of Statistics: A Concise Course in Statistical Inference*, whose exercises this sheet was inspired by.