Solution to Problem Set 1 February 18, 2025

Exercise 1.

Let us recall that a random variable $\varphi(X)$ is the conditional expectation of Y given X, denoted $\mathbb{E}[Y|X]$, if and only if it satisfies

$$\mathbb{E}[\varphi(X)h(X)] = \mathbb{E}[Yh(X)]$$

for any bounded function $h : \mathbb{R} \to \mathbb{R}$.

(a) Note that $\alpha Y + \beta Z$ is an integrable random variable. By the definition of conditional expectation, the right-hand side is a linear combination of two functions of X, and hence it is itself a function of X. Using the linearity of expectation, we have

$$\begin{aligned} \mathbb{E}\big[(\alpha\,\mathbb{E}[Y|X] + \beta\,\mathbb{E}[Z|X])h(X)\big] &= \alpha\,\mathbb{E}\big[\mathbb{E}[Y|X]h(X)\big] + \beta\,\mathbb{E}\big[\mathbb{E}[Z|X]h(X)\big] \\ &= \alpha\,\mathbb{E}\big[Yh(X)\big] + \beta\,\mathbb{E}\big[Zh(X)\big] \\ &= \mathbb{E}\big[(\alpha Y + \beta Z)h(X)\big] \end{aligned}$$

for any bounded function $h : \mathbb{R} \to \mathbb{R}$ and any $\alpha, \beta \in \mathbb{R}$. Thus,

$$\mathbb{E}[\alpha Y + \beta Z|X] = \alpha\mathbb{E}[Y|X] + \beta\mathbb{E}[Z|X].$$

(b) To prove this property, we use the following definition of $\varphi(X)$; in the discrete case:

$$\varphi(x) = \mathbb{E}[Y|X = x] = \sum_y y\mathbb{P}(Y = y|X = x),$$

for all x, and in the continuous case:

$$\varphi(x) = \mathbb{E}[Y|X = x] = \int_{\mathbb{R}} y \frac{f_{X,Y}(x, y)}{f_X(x)}\, dy,$$

where (X, Y) have a joint density function $f_{X,Y}(x, y)$.

1. <u>Discrete case</u>: Suppose first that $Y \geqslant 0$. Then,

$$\varphi(x) = \mathbb{E}[Y|X = x] = \sum_y y\mathbb{P}(Y = y|X = x) \geqslant 0$$

for all x. Thus,

$$\varphi(X) = \mathbb{E}[Y|X] \geqslant 0.$$

Considering $Z - Y \geqslant 0$ and using property (a), we obtain (b).

2. <u>Continuous case</u>: In the same way as in the discrete case, we first consider a random variable $Y \geqslant 0$. Since $f_{X,Y}(x,y) = 0$ for all $y \leqslant 0$, we have

$$\varphi(x) = \mathbb{E}[Y|X = x] = \int_{\mathbb{R}} y \frac{f_{X,Y}(x,y)}{f_X(x)}\, dy \geqslant 0$$

for all x. Thus, $\varphi(X) = \mathbb{E}[Y|X] \geqslant 0$. Property (b) follows from the above and property (a).

(c) By definition, $\mathbb{E}[Z|X,Y]$ is an integrable random variable. It remains to show that for any bounded function $h : \mathbb{R} \to \mathbb{R}$,

$$\mathbb{E}\big[\mathbb{E}[Z|X]h(X)\big] = \mathbb{E}\big[\mathbb{E}[Z|X,Y]h(X)\big].$$

The left-hand side equals $\mathbb{E}[Zh(X)]$ by definition. Since h can be viewed as a function $h : \mathbb{R}^2 \to \mathbb{R}$ with $(x,y) \mapsto h(x)$, the right-hand side is also equal to $\mathbb{E}[Zh(X)]$. This proves the result.

(d) Since f is bounded and $\mathbb{E}[|Y|] < \infty$, we have that $\mathbb{E}[|Yf(X)|] < \infty$. For any bounded function $h : \mathbb{R} \to \mathbb{R}$, we can write

$$\begin{aligned} \mathbb{E}[Yf(X)h(X)] &= \mathbb{E}[Y(f\cdot h)(X)] = \mathbb{E}\big[\mathbb{E}(Y|X)(f\cdot h)(X)\big] \\ &= \mathbb{E}\big[\mathbb{E}(Y|X)f(X)h(X)\big] \end{aligned}$$

where we use the fact that the product of two bounded functions f and h is a bounded function. Thus,

$$\mathbb{E}[Yf(X)|X] = \mathbb{E}[Y|X]f(X).$$

(e) If X and Y are independent, then $h(X)$ and Y are also independent, and the expectation of their product is the product of their expectations. Hence,

$$\mathbb{E}[Yh(X)] = \mathbb{E}[Y]\mathbb{E}[h(X)] = \mathbb{E}[\mathbb{E}[Y]h(X)]$$

for any bounded function $h : \mathbb{R} \to \mathbb{R}$. Thus, $\mathbb{E}[Y|X] = \mathbb{E}[Y]$.

(f) This is a special case of point (d) with $Y = 1$ identically.

(g) We begin by recalling the following result. If $g : \mathbb{R} \to \mathbb{R}$ is convex and $c \in \mathbb{R}$ is fixed, then there exists an affine function $l(x) = ax + b$ such that $l(c) = g(c)$ and $l(x) \leqslant g(x)$ for all x.

Using this, we can show that for a convex function $g : \mathbb{R} \to \mathbb{R}$, there exist two sequences $(a_n)_{n\in\mathbb{N}}$ and $(b_n)_{n\in\mathbb{N}}$ such that for all $x \in \mathbb{R}$, $g(x) = \sup_{n\in\mathbb{N}}(a_n x + b_n)$. Applying this to conditional expectation:

$$\mathbb{E}[g(Y)|X] \geqslant \mathbb{E}[a_n Y + b_n|X] = a_n\mathbb{E}[Y|X] + b_n, \qquad \forall n \in \mathbb{N}.$$

Taking the supremum over n yields:

$$\mathbb{E}[g(Y)|X] \geqslant \sup_{n\in\mathbb{N}}(a_n\mathbb{E}[Y|X] + b_n) = g(\mathbb{E}[Y|X]).$$

This proves (g).

Exercise 2.

Notice that the events $\{N > n\}$ and $\{X_1 + \cdots + X_n \leqslant x\}$ are the same. Thus,

$$\mathbb{P}(N > n) = \mathbb{P}(X_1 + \cdots + X_n \leqslant x).$$

For $x \in]0, 1[$, define $G_n(x) := \mathbb{P}\{X_1 + \cdots + X_n \leqslant x\}$. We will prove by induction that $G_n(x) = \frac{x^n}{n!}$.

The property holds for $n = 1$: $G_1(x) = \mathbb{P}\{X_1 \leq x\} = x$. Now assume the property holds at rank $n-1$ and prove it at rank n.

We will condition the desired probability on X_1. The definition of conditional expectation with $h(X_1) = 1$ identically gives:

$$\mathbb{P}(N > n) = \mathbb{E}[\mathbb{1}_{\{N>n\}}] = \mathbb{E}[\mathbb{E}[\mathbb{1}_{\{N>n\}}|X_1]] = \mathbb{E}[\mathbb{P}(N > n|X_1)].$$

If $f_{X_1}(u) = \mathbb{1}_{[0,1]}(u)$ is the density of X_1, then:

$$\begin{aligned}\mathbb{P}(N > n) &= \int_{\mathbb{R}} \mathbb{P}(X_1 + \cdots + X_n \leqslant x|X_1 = u) f_{X_1}(u)\, du \\ &= \int_0^1 \mathbb{P}(X_1 + \cdots + X_n \leqslant x|X_1 = u)\, du \\ &= \int_0^x \mathbb{P}(X_1 + \cdots + X_n \leqslant x|X_1 = u)\, du\end{aligned}$$

because if $u > x$, then the conditional probability is necessarily zero. Indeed, it is impossible to have $X_1 + \cdots + X_n \leqslant x$ if $X_1 > x$. Using the independence of X_1 from $X_2, \ldots, X_n$, the substitution $x - u = x_1$, and the fact that $X_1, \ldots, X_n$ are independent and identically distributed, we obtain:

$$\begin{aligned}\mathbb{P}(N > n) &= \int_0^x \mathbb{P}(X_2 + \cdots + X_n \leqslant x - u)\, du \\ &= \int_0^x \mathbb{P}(X_2 + \cdots + X_n \leqslant x_1)\, dx_1 \\ &= \int_0^x G_{n-1}(x_1)\, dx_1.\end{aligned}$$

Next, applying the induction hypothesis:

$$\begin{aligned}\mathbb{P}(N > n) &= \int_0^x G_{n-1}(x_1)\, dx_1 \\ &= \int_0^x \frac{x_1^{n-1}}{(n-1)!}\, dx_1 \\ &= \frac{x^n}{n!}.\end{aligned}$$

Exercise 3.

Note that $\mathbb{E}|X_1X_2| = \mathbb{E}|X_1|\mathbb{E}|X_2| < \infty$. Furthermore, by an application of exercise 1 we have that

$$\begin{aligned}\mathbb{E}[X_1X_2|X_3] &\overset{c)}{=} \mathbb{E}[\mathbb{E}[X_1X_2|X_2, X_3]|X_3]\\ &\overset{d)}{=} \mathbb{E}[X_2\mathbb{E}[X_1|X_2, X_3]|X_3]\\ &\overset{e)}{=} \mathbb{E}[X_2\mathbb{E}[X_1]|X_3]\\ &\overset{a)}{=} \mathbb{E}[X_1]\mathbb{E}[X_2|X_3].\end{aligned}$$

Exercise 4.

To understand where the formula comes from, let's consider the following heuristics: let $y_0 \in \mathbb{R}$ and $\varepsilon > 0$, suppose that the density h is continuous, then

$$\begin{aligned}\mathbb{P}[X \in A \mid Y \in (y_0 - \varepsilon; y_0 + \varepsilon)] &= \frac{\mathbb{P}[X \in A \,\cap\, Y \in (y_0 - \varepsilon; y_0 + \varepsilon)]}{\mathbb{P}[Y \in (y_0 - \varepsilon; y_0 + \varepsilon)]}\\ &= \frac{\int_{y_0-\varepsilon}^{y_0+\varepsilon}\int_A h(x,y)dxdy}{\int_{y_0-\varepsilon}^{y_0+\varepsilon}\int_{\mathbb{R}} h(x,y)dxdy}\\ &\overset{\varepsilon\to 0}{\approx} \frac{2\varepsilon\int_A h(x,y_0)dx}{2\varepsilon\int_{\mathbb{R}} h(x,y_0)dx} = \int_A \frac{h(x,y_0)}{\int_{\mathbb{R}} h(x',y_0)dx'}dx\end{aligned}$$

Thus, we can infer that the conditional law of X knowing $\{Y = y_0\}$ has density

$$\frac{h(\cdot, y_0)}{\int_{\mathbb{R}} h(x', y_0)dx'},$$

and thus we would expect the density formula.

Moreover, it remains to prove the condition in the definition of conditional expectation. Let $f : \mathbb{R} \to \mathbb{R}$ be bounded and let $g(Y) := \frac{\int_{\mathbb{R}} x\, h(x,Y)dx}{\int_{\mathbb{R}} h(x,Y)dx} = \frac{\int_{\mathbb{R}} x\, h(x,Y)dx}{f_Y(Y)}$ for simplicity. We will check that $\mathbb{E}[g(Y)f(Y)] = \mathbb{E}[Xf(Y)]$. In fact,

$$\begin{aligned}\mathbb{E}[g(Y)f(Y)] &= \int_{\mathbb{R}} g(y)f(y)f_Y(y)dy = \int_{\mathbb{R}}(\int_{\mathbb{R}} x\, h(x,y)dx)f(y)dy\\ &= \int_{\mathbb{R}}\int_{\mathbb{R}} xh(x,y)f(y)dxdy = \mathbb{E}[Xf(Y)].\end{aligned}$$

The first equality of the second line is followed by Fubini-Tonelli theorem.

Exercise 5.

For the first part, observe that by Exercise 1

$$\mathbb{E}\Big[\sum_{i=1}^n X_i\Big|X_1\Big] = \mathbb{E}[X_1|X_1] + \sum_{i=2}^n \mathbb{E}[X_i|X_1] = X_1 + \sum_{i=2}^n \mathbb{E}[X_i] = X_1 + (n-1)\mathbb{E}[X_1].$$

For the second question, we start by proving the following symmetry property:

Lemma 1. *Let X, Y, Z be three random variables with a finite first moment such that (X, Z) has the same law as (Y, Z). Show that,*

$$\mathbb{E}[X|Z] = \mathbb{E}[Y|Z] \quad a.s..$$

Proof. For every bounded function g the equality of the joint laws of (X, Z) and (Y, Z), resp., implies

$$\mathbb{E}[Xg(Z)] = \mathbb{E}[Yg(Z)] \quad \Rightarrow \quad \mathbb{E}\big[\mathbb{E}[X|Z]g(Z)\big] = \mathbb{E}[Xg(Z)] = \mathbb{E}[Yg(Z)].$$

This, in turn, implies that a.s. $\mathbb{E}[X|Z] = \mathbb{E}[Y|Z]$. □

Now, by lemma, for every $1 \leqslant i, j \leqslant n$ one has,

$$\mathbb{E}\Big[X_i \Big| \sum_{k=1}^{n} X_k\Big] = \mathbb{E}\Big[X_j \Big| \sum_{i=k}^{n} X_k\Big], \text{ a.s.}$$

and almost surely,

$$\sum_{i=1}^{n} \mathbb{E}\Big[X_i \Big| \sum_{k=1}^{n} X_k\Big] = \mathbb{E}\Big[\sum_{i=1}^{n} X_i \Big| \sum_{k=1}^{n} X_k\Big] = \sum_{k=1}^{n} X_k.$$

The result follows.