Probability 1 - Ex. 1 Solution

1. $X \sim Bin(n, p)$ $\quad Y \sim Ber(X/n)$

a) Describe the law of Y

Y is a r.v. supported in $\{0,1\}$. It suffices to compute $\mathbb{P}(Y=1)$:

$$\mathbb{P}(Y=1) = \sum_{j=0}^{n} \mathbb{P}(Y=1 \mid X=j)\, \mathbb{P}(X=j) = \sum_{j=0}^{n} \frac{j}{n} \binom{n}{j} p^j (1-p)^{n-j}$$

$$= \sum_{j=1}^{n} \binom{n-1}{j-1} p^{j-1} (1-p)^{(n-1)-(j-1)} \cdot p =$$

for $j>0$ $\frac{n}{j}\binom{n-1}{j-1} = \binom{n}{j}$

$$= p \underbrace{\sum_{i=0}^{n-1} \binom{n-1}{i} p^i (1-p)^{(n-1)-i}}_{(p+(1-p))^{n-1} = 1^{n-1} = 1} = p$$

So $\mathbb{P}(Y=1) = p$, $\quad \mathbb{P}(Y=0) = 1-p$ and

$$Y \sim Ber(p)$$

b) $$\mathbb{E}[X \mid Y=1] = \sum_{j=0}^{n} \mathbb{P}(X=j \mid Y=1) \cdot j =$$

$$= \sum_{j=0}^{n} \mathbb{P}(Y=1 \mid X=j) \frac{\mathbb{P}(X=j)}{\mathbb{P}(Y=1)} \cdot j = \sum_{j=0}^{n} \frac{j}{n} \frac{\binom{n}{j} p^j (1-p)^{n-j}}{p} j$$

$\mathbb{P}(A|B) = \mathbb{P}(B|A) \frac{\mathbb{P}(A)}{\mathbb{P}(B)}$

$$= \sum_{i=0}^{n} \binom{n-1}{i} p^i (1-p)^{(n-1)-(i-1)} (i+1) \quad (*)$$

Let $Z \sim Bin(n-1, p)$. Then $\mathbb{E}[Z+1] = (*)$.

On the other hand, $\mathbb{E}[Z+1] = (n-1)p + 1$.

So $\mathbb{E}[X \mid Y=1] = (n-1)p + 1$

Recall that $\mathbb{E}[\mathbb{E}[X \mid Y]] = \mathbb{E}[X]$

$\Rightarrow \mathbb{E}[X \mid Y=0]\,\mathbb{P}(Y=0) + \mathbb{E}[X \mid Y=1]\,\mathbb{P}(Y=1) = \mathbb{E}[X]$.

Since $\mathbb{E}[X \mid Y=1] = np+1$, $\mathbb{P}(Y=1) = p$, $\mathbb{P}(Y=0) = 1-p$ and $\mathbb{E}[X] = np$, we have

$$\mathbb{E}[X \mid Y=0] = \frac{np - p((n-1)p+1)}{1-p} = \frac{np - np^2 - p + p^2}{1-p} = np - p$$

$$\mathbb{E}[X \mid Y=0] = (n-1)p$$

Conclusion: $\mathbb{E}[X \mid Y] = (n-1)p + Y$

---

(2) $X, Y \sim \mathrm{Unif}(\{1, \dots, 6\})$ $\quad S = X + Y$

$X \perp\!\!\!\perp Y$

$\mathbb{E}[\cdot \mid \cdot]$	X	Y	S
X	X (2)	$7/2$ (1)	$X + 7/2$ (3)
Y	$7/2$ (1)	Y (2)	$7/2 + Y$ (3)
S	$\frac{1}{2}S$ (4)	$\frac{1}{2}S$ (4)	S (2)

Obs 1: If $A \perp\!\!\!\perp B$ are discrete r.v., then $\mathbb{E}[A \mid B] = \mathbb{E}[A]$

Proof: $\mathbb{E}[A \mid B=j] = \sum_{i \in \mathrm{Range}(A)} \mathbb{P}(A=i \mid B=j) \cdot i$

$\mathbb{P}(A=i \mid B=j) = \mathbb{P}(A=i)$ by $A \perp\!\!\!\perp B$

$= \sum_{i \in \mathrm{Range}(A)} \mathbb{P}(A=i)\, i = \mathbb{E}[A]$

Obs 2: A discrete r.v. Then $\mathbb{E}[A \mid A] = A$

Proof: $\mathbb{E}[A \mid A=j] = \sum_{i \in Range(A)} \mathbb{P}(A=i \mid A=j) \cdot i = \mathbb{P}(A=j \mid A=j) \cdot j = j$

Obs 3: If X, Y_1, Y_2 are discrete r.v. then
$$\mathbb{E}[aY_1 + bY_2 \mid X] = a\,\mathbb{E}[Y_1 \mid X] + b\,\mathbb{E}[Y_2 \mid X]$$

Proof: $a\,\mathbb{E}[Y_1 \mid X=j] + b\,\mathbb{E}[Y_2 \mid X=j] =$

$$a \sum_{i_1 \in Range(Y_1)} i_1\, \mathbb{P}[Y_1 = i_1 \mid X=j] + b \sum_{i_2 \in Range(Y_2)} i_2\, \mathbb{P}(Y_2 = i_2 \mid X=j)$$

$$= \sum_{\substack{i_1 \in Range(Y_1) \\ i_2 \in Range(Y_2)}} (a i_1 + b i_2)\, \mathbb{P}(Y_1 = i_1, Y_2 = i_2 \mid X=j)$$

$$= \sum_{i \in Range(aY_1 + bY_2)} i \sum_{\substack{i_1 \in Range(Y_1) \\ i_2 \in Range(Y_2) \\ a i_1 + b i_2 = i}} \mathbb{P}(Y_1 = i_1, Y_2 = i_2 \mid X=j)$$

$$= \sum_{i \in Range(aY_1 + bY_2)} i \quad \mathbb{P}(aY_1 + bY_2 = i \mid X=j) = \mathbb{E}[aY_1 + bY_2 \mid X=j]$$

It follows from obs 3 that $\mathbb{E}[S \mid X] = \mathbb{E}[X \mid X] + \mathbb{E}[Y \mid X]$
$$= X + 7/2$$

Obs 4: If X, Y, S are as given in Ex2, and if

$\mathbb{P}(S=i) \neq 0$

$$\mathbb{P}(X-Y=j \mid S=i) = \frac{\mathbb{P}\left(X=\frac{i+j}{2}, Y=\frac{i-j}{2}\right)}{\mathbb{P}(S=i)} \overset{X \perp\!\!\!\perp Y}{=} \frac{\mathbb{P}\left(X=\frac{i+j}{2}\right)\mathbb{P}\left(Y=\frac{i-j}{2}\right)}{\mathbb{P}(S=i)}$$

$$\underset{X,Y \text{ have the same law}}{=} \frac{\mathbb{P}\left(X=\frac{i-j}{2}\right)\mathbb{P}\left(Y=\frac{i+j}{2}\right)}{\mathbb{P}(S=i)} = \dots = \mathbb{P}(X-Y=-j \mid S=i)$$

obs 3

It follows that $\mathbb{E}[X-Y \mid S=i]=0$, so $\mathbb{E}[X \mid S]=\mathbb{E}[Y \mid S]$

But $\mathbb{E}[X \mid S]+\mathbb{E}[Y \mid S] \overset{\text{obs(3)}}{=} \mathbb{E}[X+Y \mid S] \overset{\text{obs(2)}}{=} S$

So $\mathbb{E}[X \mid S]=\mathbb{E}[Y \mid S]=\frac{1}{2} S$

---

3 a) If X is a step function,

then $\operatorname{im} X$ is finite (in fact, it has up to 2^k many elements)

Let $\operatorname{im} X=\{a_1, \dots, a_j\}$ and define $A_i=X^{-1}(\{a_i\}) \in \sigma(Y)$, because X is $\sigma(Y)$-measurable.

Clearly $X=\sum_{i=1}^{l} a_i \mathbb{1}_{A_i}$, and $A_i \neq \emptyset$ for any i.

By definition of $\sigma(Y)$, there are $B_1, \dots, B_j\}$ Borel subsets of $\mathbb{R}$ s.t. $A_i = Y^{-1} B_i)$.

Because the $\{A_i\}_i$ are disjoint, we have that the $B_i\}_i$ are also disjoint, as $B_i \cap B_j \subseteq Y(A_i \cap A_j)=\emptyset$ for $i \neq j$.

We construct $f: \mathbb{R} \to \mathbb{R}$ explicitly:

Define $f(B_i):=a_i$, $f(x)=0$ for $x \in \bigcap_{i=1}^{l} B_i^C$ also a Borel set!!

Because each B_i is a Borel set f is measurable.

Because the family $\{B_i\}_i$ is a disjoint cover of $\mathbb{R}$, f is unique and well defined.

Finally, if $X(\omega)=a_i$ $\omega \in A_i$ so $f(Y(\omega)) \in f(B_i)=a_i$ so $X=f(Y)$ ☒

(b) Let X_n be a step function approximation of X s.t. $X_n \to X$ pointwise (sure convergence)

For instance, we can take

$$X_n := \lfloor 2^n \min(X, n) \rfloor \frac{1}{2^n}$$ We can directly observe that $X_n \nearrow X$ pointwise, that is

- $X_n(\omega) \le X_{n+1}(\omega) \quad \forall n \ge 0 \quad \omega \in \Omega$
- $\lim_{n \to +\infty} X_n(\omega) = X$

Also, X_n is clearly a step function, as it only takes values in $\{ \frac{k}{2^n} \mid k = 0, 1, \dots, 2^n \times n \}$.

Finally, X_n is $\sigma(Y)$-measurable because

$$\{ X_n = \frac{k}{2^n} \} = \{ X - [\frac{k}{2^n}, \frac{k+1}{2^n}) \} \quad \text{for } k < 2^n \times n$$

$$\{ X_n = n \} = \{ X \in [n, +\infty) \} \quad \longrightarrow \in \sigma(Y)$$

because X as $\sigma(Y)$-measurable.

Let $X_n = f_n(Y)$. Because $X_n \nearrow X$, then $f_n(x) \le f_{n+1}(x)$ for any $x \in \operatorname{im} Y$. This defines $f = \lim_n f_n$ Borel-measurable such that

$$f(Y) = X \quad \square$$

(c) Let $X^+ = \max(0, X)$ $\sigma(Y)$-measurable funcs.

$X^- = \max(0, -X)$

$X = X^+ - X^-$

Then $X^+, X^- \geq 0$ so there is f^+, f^- s.t.

$$f^+(Y) = X^+ , \quad f^-(Y) = X^-.$$

if $f = f^+ - f^-$ we have $f(Y) = X$. It suffices to argue that f is Borel-measurable.

For that, it is enough to show that $f^{-1}(a, +\infty) \in \mathcal{B}(\mathbb{R})$ for $a \geq 0$, and that $f^{-1}(-\infty, b) \in \mathcal{B}(\mathbb{R})$ for $b \leq 0$.

Note that $f^{-1}(a, +\infty) = (f^+)^{-1}(a, +\infty) \in \mathcal{B}(\mathbb{R})$ because f^+, f^- are

$$f^{-1}(-\infty, b) = (f^-)^{-1}(-\infty, b) \in \mathcal{B}(\mathbb{R})$$

measurable by hypothesis.

Ex 4

(a) x satisfies both $x \in \mathcal{L}$
$\|x - x\| = 0.$

Further, if any $y \in \mathcal{L}$ is s.t. $\|y - x\| \leq 0$, then by the axioms of the norm, $y - x = 0$, so $y = x$.
Then $x = \Pi x$

(b) From (a) and because $\Pi x \in \mathcal{L}$, $\Pi^2 x = \Pi x$.

(c) **Part 1:** $\forall y \in \mathcal{L}, \quad \langle y, \Pi x \rangle = \langle y, x \rangle \quad (*)$

Let $y \in \mathcal{L}$ be generic, and consider $\alpha \in \mathbb{R}$. We have

$$\| (\Pi x - \alpha y) - x \|^2 \geq \| \Pi x - x \|^2 \qquad (1)$$

But $\|(\Pi x - \alpha y) - x\|^2 = \langle (\Pi x - x) - \alpha y, (\Pi x - x) - \alpha y \rangle$

$$= \langle \Pi x - x, \Pi x - x \rangle - 2\alpha \langle \Pi x - x, y \rangle + \alpha^2 \langle y, y \rangle$$

Using this in (1) gives us

$$\cancel{\langle \Pi x - x, \Pi x - x \rangle} - 2\alpha \langle \Pi x - x, y \rangle + \alpha^2 \langle y, y \rangle \geq \cancel{\|\Pi x - x\|^2}$$

$$\Rightarrow \alpha^2 \langle y, y \rangle - 2\alpha \langle \Pi x - x, y \rangle \geq 0 \qquad (2)$$

So either $y = 0$ or this is a quadratic inequality

If $y = 0$, $(*)$ trivially holds

If $y \neq 0$, take $\alpha = \dfrac{\langle \Pi x - x, y \rangle}{\langle y, y \rangle}$ in (2) to get

$$\frac{\langle \Pi x - x, y \rangle^2}{\langle y, y \rangle} - 2 \frac{\langle \Pi x - x, y \rangle^2}{\langle y, y \rangle} \geq 0$$

$$\Rightarrow \frac{\langle \Pi x - x, y \rangle^2}{\langle y, y \rangle} \leq 0 \Rightarrow \langle \Pi x - x, y \rangle = 0$$

which is equivalent to $(*)$ $\square$

Rem: It follows that $\langle \Pi x - x, y \rangle = 0 \;\; \forall y \in L$, recovering Prop 2.7

Part 2 Uniqueness

Let $a \in L$ such that $\forall y \in L$, $\langle a, y \rangle = \langle x, y \rangle$.

Then $\langle a, y \rangle = \langle \Pi x, y \rangle \;\; \forall y \in L$, by the above.

Pick $y = \Pi x - a$ to obtain

$\langle a, \Pi x - a \rangle = \langle \Pi x, \Pi x - a \rangle \Leftrightarrow \langle \Pi x - a, \Pi x - a \rangle = 0$

$\Leftrightarrow \| \Pi x - a \|^2 = 0 \Leftrightarrow \Pi x = a$

Concluding the uniqueness. $\square$

d) If $x \in \mathcal{L}^\perp$, by c) we have that

$\|\Pi x\|^2 = \langle \Pi x, \Pi x \rangle = \langle \Pi x, x \rangle \underset{x \in \mathcal{L}^\perp}{=} 0$

So $\Pi x = 0$ $\square$

Last modified: 13:54