Stat 2911 Lecture Notes
Class 8, 2017

Uri Keich

Coupon collector problem (Rice 4.1), expectation of a product of ind RVs, L^2 is a vector subspace of L^1 corresponding to RVs with finite variance

Fubini's Theorem

(a) A positive series can be summed up in any order.

(b) An absolutely convergent series can be summed up in any order.

Let $g: \mathbb{R}^n \to \mathbb{R}$ and let $Z = g(\underline{X})$ where $\underline{X}$ is an n-dimensional random vector.

Claim.

(i) $Z \in L^1$ iff $\sum_{\underline{x} \in \mathbb{R}^n} |g(\underline{x})| \, p_{\underline{X}}(\underline{x}) < \infty$

(ii) If $Z \in L^1$ then $E(Z) = \sum_{\substack{\underline{x} \in \mathbb{R}^n: \\ p_{\underline{X}}(\underline{x}) > 0}} g(\underline{x}) \, p_{\underline{X}}(\underline{x})$
(or $Z \geq 0$, or $Z \leq 0$)

Cor. If X & Y are jointly distributed RVs with $X, Y \in L^1$ and $\alpha, \beta \in \mathbb{R}$ then

(i) $\alpha X + \beta Y \in L^1$ (L^1 is a vector space)

(ii) $E(\alpha X + \beta Y) = \alpha E(X) + \beta E(Y)$

(expectation is linear <u>regardless</u> of dependency)

Note. If we take $Y \equiv 1$ (constant RV) then

$$E(\alpha X + \beta) = \alpha E(X) + \beta$$

By induction, if $X_1, \ldots, X_n \in \mathcal{L}^1(\Omega)$ then $\forall \alpha_i \in \mathbb{R}$, $\sum_1^n \alpha_i X_i \in \mathcal{L}^1$ and $E\left(\sum_1^n \alpha_i X_i\right) = \sum_1^n \alpha_i E(X_i)$

Examples

1) $X_i \sim \text{Bernoulli}(p)$ RVs (not necessarily ind.)

Let $S_n = \sum_1^n X_i$

Then,

$$
\begin{aligned}
E(S_n) &= E\left(\sum_1^n X_i\right) \\
&= \sum_1^n E(X_i) \\
&= \sum_1^n p \\
&= np
\end{aligned}
$$

In particular this shows that if $Y \sim \text{Binom}(n,p)$ then $E(Y) = np$.

(The X_i are ind. in this case)

2) Coupon Collector Problem

n types of coupon.
At each trial a single randomly selected coupon is collected.
What is the expected number of trials for completing a collection with all n types?

Let S_n = number of trials for collecting all n types

X_1 = # of trials for accumulating 1 type
X_2 = # of additional trials for owning 2 types
X_3 = " " " 3 "
⋮
X_n = " " " n "

$$S_n = \sum_{i=1}^{n} X_i \implies E(S_n) = \sum_{i=1}^{n} E(X_i)$$

$$X_1 = 1 \implies E(X_1) = 1$$
$$X_2 \sim \text{geometric}\left(\frac{n-1}{n}\right) \implies E(X_2) = \frac{n}{n-1}$$
$$X_3 \sim \text{geometric}\left(\frac{n-2}{n}\right) \implies E(X_3) = \frac{n}{n-2}$$
$$\vdots \qquad\qquad \vdots$$
$$X_i \sim \text{geometric}\left(\frac{n-(i-1)}{n}\right) \implies E(X_i) = \frac{n}{n-i+1}$$

$$\implies E(S_n) = \sum_{i=1}^{n} \frac{n}{n-i+1} = n\left(\frac{1}{n} + \frac{1}{n-1} + \dots + \frac{1}{2} + 1\right)$$

$$\sim n \log n$$

Aside: Are X_i ind RVs? Does it matter?
Much simpler than finding the pmf of S_n

$E(X+Y)=E(X)+E(Y)$ but is $E(XY)=E(X)E(Y)$?

Example. Let $X=Y\sim \text{Bernoulli}(p)$, then

$$E(XY)=E(X^2)=p\neq p^2=E(X)E(Y)$$

However, if X,Y are ind Bernoulli(p), then

$$E(XY)=1\cdot p^2=E(X)E(Y)$$

Is this the general rule?

Moreover, do we even know that $X\cdot Y\in L^1$?

Generally, $X,Y\in L^1 \not\Rightarrow X\cdot Y\in L^1$ (later)

Claim. Suppose $X,Y\in L^1$ are ind then $XY\in L^1$ and

$$E(XY)=E(X)E(Y)$$

Proof. Let $g:\mathbb{R}^2\to\mathbb{R}$ be $g(x,y)=x\cdot y$ and define $Z=g(X,Y)=X\cdot Y$.

$$\sum_{x,y}{}' |g(x,y)|\,p_{XY}(x,y)=\sum_{x,y}{}' |xy|\,p_X(x)\,p_Y(y)$$

$$\overset{?}{=}\sum_x{}' |x|\,p_X(x)\left(\sum_y{}' |y|\,p_Y(y)\right)\overset{?}{<}\infty$$

$$\overset{?}{\Rightarrow} Z=XY\in L^1 \text{ and } E(Z)\overset{?}{=}E(X)E(Y)$$

Repeat the above argument w/o the $|\cdot|$ and use Fubini (ii)

Recall:

$L^1 = \{ \underline{X} : \sum_1 |x_k| \, p_x(x_k) < \infty \}$ $\quad [\underline{X} \in L^1 \overset{?}{\Longleftrightarrow} E|\underline{X}| < \infty]$

$L^2 = \{ \underline{X} : \sum_1 x_k^2 \, p_x(x_k) < \infty \}$ $\quad [\underline{X} \in L^2 \Longleftrightarrow E\underline{X}^2 < \infty$

$\overset{?}{\Longleftrightarrow} \underline{X}^2 \in L^1]$

What is the relation between L^1 and L^2?

Let $\underline{X}$ be a RV with pmf $p_x(n) = c/n^3 \cdot 1_{n \in \mathbb{N}}$, where $c = 1/\sum_1^\infty n^{-3}$.

Is $\underline{X} \in L^1$?

$\sum_1^\infty n p_x(n) = \sum_1^\infty c/n^2 \qquad \infty \Rightarrow \underline{X} \qquad L^1$

Is $\underline{X} \in L^2$?

$\sum_1^\infty n^2 p_x(n) = \sum_1^\infty c/n \qquad \infty \Rightarrow \underline{X} \qquad L^2$

An example of $\underline{X} \in L^1 \setminus L^2$ or $\underline{X} \in L^1 \not\Rightarrow \underline{X} \in L^2$.

What about the converse?

Claim. $\underline{X} \in L^2 \Rightarrow \underline{X} \in L^1$

Proof. $|x| < 1 + x^2 \quad \forall x \in \mathbb{R}$ (exercise)

$$\Rightarrow \sum_k |x_k| \, p_x(x_k) < \sum (1 + x_k^2) \, p_x(x_k)$$

$$= \sum p_x(x_k) + \sum x_k^2 p_x(x_k) < \infty \qquad \square$$

Recall: $X \in L^1$ iff $E(X)$ exists and is finite, we now have

Cor. $X \in L^2$ iff $V(X)$ exists and is finite

Proof. $X \in L^2 \Rightarrow X \in L^1 \Rightarrow E(X) \in \mathbb{R}$

$$\Rightarrow V(X) = E(X^2) - E^2(X) \in \mathbb{R}$$

Conversely, we showed that, if $V(X)$ exists and is finite then $E(X^2) < \infty$.

Suppose X and Y are jointly distributed RVs with finite variance.

Want $V(X+Y)$ but is $X+Y \in L^2$?

The answer is positive but to see that we first need to verify that $X \cdot Y \in L^1$.

Aside. Recall we said $X, Y \in L^1 \not\Rightarrow X \cdot Y \in L^1$

Can you think of an example now?

Take $X = Y$ the same RV with $p_x = \frac{c}{n^3} 1_{n \in \mathbb{N}}$.

We saw that $X \in L^1$ but $X \notin L^2$, so

$$X \cdot Y = X^2 \notin L^1$$

<u>Claim.</u> If $X, Y \in L^2$ then $XY \in L^1$

<u>Proof.</u> For $x, y \in \mathbb{R}$: $0 \leq (|x| - |y|)^2 = x^2 - 2|x||y| + y^2$

$\Rightarrow |xy| \leq x^2/2 + y^2/2 = \frac{x^2 + y^2}{2}$

geometric mean of x^2 and y^2 is $\leq$ arithmetic mean

Again, using $Z = g(X, Y) = X \cdot Y$

$$\Rightarrow \sum_{x,y}' |xy| \, p_{XY}(x,y) \leq \sum_{x,y}' \left(x^2/2 + y^2/2 \right) p_{XY}(x,y)$$

$$\overset{?}{=} \sum_x' x^2/2 \sum_y' p_{XY}(x,y) + \sum_y' y^2/2 \sum_x' p_{XY}(x,y)$$

$$= \frac{1}{2} \sum_x' x^2 p_X(x) + \frac{1}{2} \sum_y' y^2 p_Y(y) \overset{?}{<} \infty$$

$\overset{?}{\Rightarrow} XY \in L^1$ $\square$

<u>Cor.</u> If $X, Y \in L^2$ then $\forall \alpha, \beta \in \mathbb{R}$: $\alpha X + \beta Y \in L^2$

L^2 is a vector space (subspace of L^1)

<u>Proof.</u> $(\alpha X + \beta Y)^2 = \alpha^2 X^2 + 2\alpha\beta XY + \beta^2 Y^2$

$X^2, Y^2, XY \overset{?}{\in} L^1$ and the RHS is a linear comb. of these three L^1-RVs, therefore it is in L^1 itself.

$\Rightarrow \alpha X + \beta Y \in L^2$

We will soon find $V(X+Y)$ but first consider the following special case.

Claim. If $V(X) = \sigma^2 < \infty$, then $\forall \alpha, \beta \in \mathbb{R}$

$$V(\alpha X + \beta) = \alpha^2 \sigma^2$$

scale shift

(variance is shift-invariant and it scales quadratically)

Proof.

Let $\mu = E(X) \overset{?}{\in} \mathbb{R}$ then

$$E(\alpha X + \beta) = \alpha \mu + \beta$$

$$\Rightarrow \quad V(\alpha X + \beta) = E[(\alpha X + \beta) - (\alpha \mu + \beta)]^2$$

$$= E[\alpha (X - \mu)]^2$$

$$= \alpha^2 E(X - \mu)^2$$

$$= \alpha^2 V(X)$$

$\square$