Stat 2911 Lecture Notes

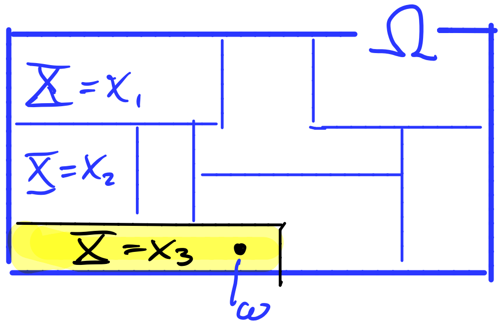
Class 19, 2017

Uri Keich

Conditional variance and the variance of the conditional expectation (Rice 4.4), Properties of CDF, Continuity of probability measure, Continuous RVs (Rice 2.2)

For $Y \in L^1$, $E(Y|X) = h(X)$, where $h(x) = E(Y|X=x)$.

For $Y \in L^2$, $V(Y|X) = g(X) - h^2(X)$, $g(x) = E(Y^2|X=x)$.



If $\omega \in \Omega$ satisfies $X(\omega) = x_3$, then

$$E(Y|X)(\omega) = E(Y|X=x_3)$$

$$V(Y|X)(\omega) = V(Y|X=x_3)$$

Thm. If $Y \in L^1$ then $E[E(Y|X)] = E(Y)$. (law of total expectation)

Random Sum: X_i are iid RVs and N is an ind. $\mathbb{N}$-valued RV. $T = \sum_1^N X_i$ is a random sum.

Claim. If $n \in N(\Omega)$, then the dist. of T given $N = n$ is the same as the dist. of $S_n = \sum_1^n X_i$.

Cor. If $X_i \in L^1$, $E(T|N=n) = n\,E(X_i)$, and if $X_i \in L^2$, $V(T|N=n) = n\,V(X_i)$

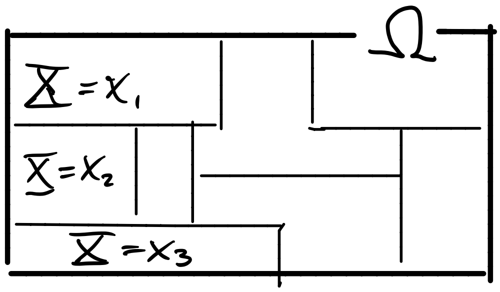
$$\Rightarrow E(T|N) \overset{X_i \in L^1}{=} N \cdot E(X_i), \quad V(T|N) \overset{X_i \in L^2}{=} N \cdot V(X_i)$$

If $X_i, N \in L^1$ then $E(T) = E[E(T|N)] = E(N)E(X_i)$.

Faulty monitor: $V(T) = \lambda p \underset{p \in (0,1)}{>} \lambda p^2 = V[E(T|N)]$

Theorem. If $y \in L^2$ then

$$V(y) = V[E(y|\underline{X})] + E[V(y|\underline{X})]$$



$V[E(y|\underline{X})]$ = the variance between the blocks/sets

$E[V(y|\underline{X})]$ = the average variance within the blocks

Proof.

$$E[V(y|\underline{X})] = E[g(\underline{X}) - h^2(\underline{X})]$$
$$= E[g(\underline{X})] - E[h^2(\underline{X})]$$

$$V[\underbrace{E(y|\underline{X})}_{h(\underline{X})}] = E[h^2(\underline{X})] - E^2[h(\underline{X})]$$

$$\Rightarrow E[V(y|\underline{X})] + V[E(y|\underline{X})]$$
$$= E[g(\underline{X})] - E^2[h(\underline{X})]$$
$$= E[E(y^2|\underline{X})] - \{E[E(y|\underline{X})]\}^2$$
$$\overset{?}{=} E(y^2) - [E(y)]^2$$
$$= V(y).$$

Back to the random sum $T = \sum_{1}^{N} X_i$, where X_i are iid L^2-RVs ind. of $N \in L^2$.

Exercise. $T \in L^2$

$$\Rightarrow \quad V(T) = V[E(T|N)] + E[V(T|N)]$$
$$= V[N \cdot E(X_i)] + E[N \cdot V(X_i)]$$
$$= E^2(X_i)\, V(N) + V(X_i)\, E(N)$$

Example. The faulty monitor: $X_i \sim \text{Bernoulli}(p)$, $N \sim \text{Poisson}(\lambda)$

$$V(T) = p^2 \cdot \lambda + p(1-p) \cdot \lambda$$
$$= \lambda p.$$

We knew that because $T \sim \text{Poisson}(\lambda p)$, but the theorem allows us to compute $V(T)$ without requiring the entire dist. of T (nor of X_i or N).

Continuous RVs

We are often interested in modeling random quantities that can take a continuous set of possible values, e.g., lifetime, tempratures, height of an individual.

The pmf cannot be used to specify such a distribution. Indeed, for such RVs we typically have $P(X=x)=0 \ \forall x$. For that we need the CDF (useful for any dist.).

Def. The cumulative distribution function (CDF) of a RV X is defined as:

$$F_X(x) = P(X \le x) .$$

Claim. If F is a CDF of a RV X with $P(X\in\mathbb{R})=1$, then:

(i) F is non-decreasing

(ii) F is right-continuous: $\lim_{x\to x_0^+} F(x) = F(x_0)$

(iii) $\lim_{x\to\infty} F(x)=1$ and $\lim_{x\to-\infty} F(x)=0$

Conversely, if F satisfies (i)-(iii) then F is the CDF of some RV X with $P(X\in\mathbb{R})=1$.

Proof. (i) $a\le b \Rightarrow \{X\le a\} \subset \{X \le b\}$

$$\Rightarrow F(a) = P(X\le a) \le P(X\le b) = F(b) .$$

To prove (ii) and (iii) we need the following lemma.

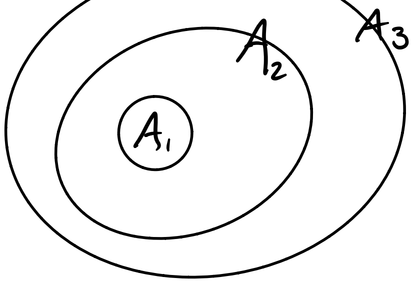
<u>Lemma</u>. (Continuity of the probability measure)

(I) If the events A_n are increasing ($A_n \uparrow$), i.e., if $A_n \subset A_{n+1}$, then $P(\bigcup_n A_n) = \lim_n P(A_n)$.

(II) If $A_n \downarrow$, i.e., if $A_n \supset A_{n+1}$, then $P(\bigcap_n A_n) = \lim_n P(A_n)$.

<u>Proof</u> (of lemma).

$A_n \uparrow$:



Define:

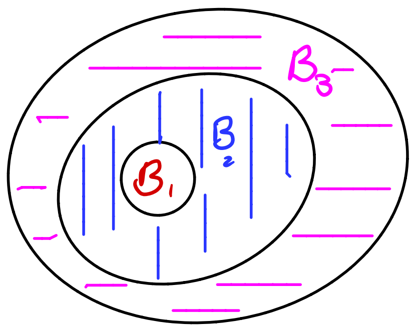
$B_1 = A_1$

$B_2 = A_2 \setminus A_1$

$B_3 = A_3 \setminus A_2$

$\vdots$

$B_n = A_n \setminus A_{n-1}$



B_n are disjoint events and $\bigcup_{k=1}^{n} B_k =$.

$$\Rightarrow \quad P\left(\bigcup_{n=1}^{\infty} A_n\right) = P\left(\bigcup_{n=1}^{\infty} B_n\right)$$

$$= \sum_1^{\infty} P(B_n)$$

$$= \lim_{N \to \infty} \sum_1^{N} P(B_n)$$

$$= \lim_{N \to \infty} P\left(\bigcup_1^{N} B_n\right)$$

$$= \lim_{N \to \infty} P(A_N)$$

$\square$

(II) Exercise. Hint: If $A_n \downarrow$ then $A_n^c \uparrow$.

Why "continuity Lemma"?

Because $A_n \uparrow$, $A_n = \bigcup_1^n A_k$ and by the lemma:

$$\lim_n P(A_n) = P\left(\bigcup_1^\infty A_n\right) = P\left(\text{"}\lim_{n\to\infty} \bigcup_1^n A_k\text{"}\right) = P\left(\text{"}\lim_n A_n\text{"}\right).$$

Analogously to $\lim_n f(x_n) = f(\lim x_n)$ for a continuous f.

Back to proving (iii) of claim.

Let $A_n = \{X \leq n\}$, then $P(A_n) = F(\quad)$.

Clearly, $A_n \quad$ so by the lemma

$$P(\cup A_n) = \lim P(A_n) = \lim_n F(n)$$

Now,

$$\cup A_n = \bigcup_{n=1}^\infty \{X \leq n\} = \{X \qquad \}$$

$$\Rightarrow \lim_n F(n) = P(\cup A_n) = P(X < \infty) =$$

Note. $\varphi(x) = \sin(2\pi x)$ satisfies $\lim_n \varphi(n) = 0$ but $\lim_{x\to\infty} \varphi(x) \neq 0$

Finally, because F is monotone $\lim_{x\to\infty} F(x)$ exists, and since $\lim_{n\to\infty} F(n) = 1$, it follows that $\lim_{x\to\infty} F(x) = 1$.

As for the second half of (iii), let $A_n = \{ X \leq -n \}$.
A_n , so
$$P(\cap A_n) = \lim_n P(A_n) = \lim_n F(\quad)$$
But
$$\bigcap_1^\infty A_n = \bigcap_1^\infty \{ X \leq -n \} = \{ X \qquad \}$$
$$\Rightarrow \lim_{n\to\infty} F(-n) = P\left(\bigcap_1^\infty A_n\right) =$$

How would you finish this proof?

(ii) F is right-continuous at x_0.
Exercise. Hint: consider $A_n = \{ X \leq x_0 + 1/n \}$.

<u>Def</u>. A RV $X : \Omega \to \mathbb{R}$ with CDF F is called a continuous RV if there exists an integrable function $f \geq 0$ s.t.
$$F(x) = \int_{-\infty}^{x} f(t)\, dt .$$
In this case, f is called the density, aka, the pdf (probability density function) of X.

<u>Comments</u>. For a cont. RV X with CDF F and pdf f.

(i) Technically, X is an absolutely continuous RV.

(ii) F is a cont. function, and by the fundamental theorem of calculus, for any x where f is cont. at x, $F'(x) = f(x)$.

(iii) If F_X, the CDF of <u>some</u> RV X, is differentiable then X is a cont. RV and $f = F_X'$ is its density. The same holds if F_X is cont. and it is differentiable everywhere except on a finite set of points.

(iv) For $a<b$,

$$\int_a^b f(t)\,dt = \int_{-\infty}^{b} f(t)\,dt - \int_{-\infty}^{a} f(t)\,dt$$

$$= F(b) - F(a)$$

(always) $$= P(X \le b) - P(X \le a)$$

(always) $$\stackrel{?}{=} P(X \in (a,b]).$$ $(= F(b)-F(a))$ always holds!

with considerable effort $\Rightarrow$ For any (measurable) set $B \subset \mathbb{R}$,

$$P(X \in B) = \int_B f(t)\,dt. \qquad (*)$$

(v) <u>Claim</u>. If X is a cont. RV then for any $x \in \mathbb{R}$,

$$P(X = x) = 0.$$

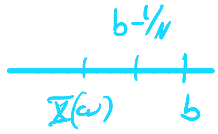
This leads to the following intriguing identity:

$$1 = P(X \in \mathbb{R}) = P\left(\bigcup_{x \in \mathbb{R}} \{X = x\}\right) = \sum_x P(X = x) = 0$$

<u>Proof</u>. You can take $B = \{x\}$ in (*) above.

Alternatively, we can prove this directly:

$$\{\omega : X(\omega) = b\} = \bigcap_n \underbrace{\{b - \tfrac{1}{n} < X \le b\}}_{B_n}$$



$$\Rightarrow P(X = b) = P\left(\bigcap_n B_n\right)$$

always holds

$$= \lim_n P(B_n)$$ since $B_n \downarrow$

$$= \lim_n P(X \in (b - 1/n, b])$$

$$\overset{?}{=} \lim_n \left[F(b) - F(b - \tfrac{1}{n})\right]$$

$$= 0$$ since F is cont. $\square$

(vii) The proof of (v) shows that for <u>any</u> RV X,

$P(X = b) = F(b) - \lim_n F(b - 1/n)$. The corollary is that the CDF always determines the pmf.