Lc 32 03/28

7.1 Properties of Expectation.

Recall: $E[X] = \sum_x x\, p(x)$ for discrete X,

$E[X] = \int_{-\infty}^{\infty} x f(x)\, dx$ } for continuous X.

$E[g(X)] = \int_{-\infty}^{\infty} g(x) f(x)\, dx$ }

Prop $\boxed{E[g(X,Y)] = \iint_{-\infty}^{\infty} g(x,y)\, f(x,y)\, dx}$

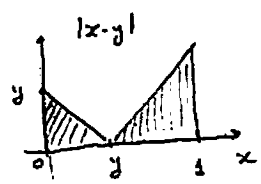
(See Ex p.68 for a similar problem)

Ex Two people agreed to meet btw 12:00 and 1:00pm. They arrive at random times, unif. distr. in this interval. Expected waiting time?

$X, Y \sim \text{Unif}(0,1)$ indep. $E[|X-Y|] = ?$

$f(x,y) = 1, \quad 0<x,y<1.$

$E[|X-Y|] = \int_0^1 \underbrace{\int_0^1 |x-y|\, dx}_{= \frac{y^2}{2} + \frac{(1-y)^2}{2}} dy$



$= \int_0^1 \frac{y^2}{2}\, dy + \int_0^1 \frac{(1-y)^2}{2}\, dy = \int_0^1 y^2\, dy = \boxed{\frac{1}{3}}.$ $= \underline{20 \text{ min}}$

(same)

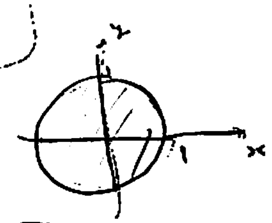
Ex! (Dart: see Ex. p.52,54 for an alternative solution based on computing the distribution of the dist.) A dart is thrown at a round board of radius 1. Compute the expected dist. to the center.

$(X,Y) \sim \text{Unif}(C)$ (circle, π)

$E\sqrt{X^2+Y^2} = ?$

$f(x,y) = \frac{1}{\pi}, \quad x^2+y^2<1$

$E\left[\sqrt{X^2+Y^2}\right] = \iint_C \sqrt{x^2+y^2}\, \frac{1}{\pi}\, dx\, dy$ $\rightarrow$ polar coord:



$= \frac{1}{\pi} \int_0^{2\pi} \int_0^1 r \cdot r\, dr\, d\theta = \frac{2\pi}{\pi} \int_0^1 r^2\, dr = 2\frac{r^3}{3} = \boxed{\frac{2}{3}}.$

• Alternative solution: By HW, #6.52, the distr. of a random pt in polar coord. is

$f_{R,\Theta}(r,\theta) = \frac{r}{\pi}, \quad 0<r<1, \quad 0<\theta<2\pi$; $2r \cdot \frac{1}{2\pi}$

$\Rightarrow f_R(r) = 2r, \quad 0<r\le 1. \quad \Rightarrow E[R] = \int_0^1 r \cdot 2r\, dr = \boxed{\frac{2}{3}}.$

THM $\forall X, Y: \quad E[X+Y] = E[X] + E[Y]$

- Not even independent!
- We formulated and used this before, but never proved.

(for continuous).

$$E[X+Y] = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x+y)\, f(x,y)\, dx\,dy$$

$$= \int_{-\infty}^{\infty} x\left(\int_{-\infty}^{\infty} f(x,y)\,dy\right) dx + \int_{-\infty}^{\infty} y\left(\int_{-\infty}^{\infty} f(x,y)\,dx\right) dy$$

$$= \int_{-\infty}^{\infty} x f_X(x)\,dx + \int_{-\infty}^{\infty} y f_Y(y)\,dy = E[X] + E[Y].$$

COR $E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i]$ (by induction)

→ Recall hat problem (Ex. 2h)

PROP $X \geq 0$ (always) $\Rightarrow E[X] \geq 0$

(for cont.)

$$E[X] = \int_{0}^{\infty} \underbrace{x}_{\geq 0}\, \underbrace{f(x)}_{\geq 0}\, dx \geq 0.$$

COR $X \geq Y \Rightarrow E[X] \geq E[Y]$.

$X - Y \geq 0 \Rightarrow E[X-Y] \geq 0$

$E[X-Y] = E[X] - E[Y]$.

COR $a \leq X \leq b \Rightarrow a \leq E[X] \leq b$

(follows from previous cor)

SKIPPED in 2011

~~Lec 30 | 03/?/?~~

7.2 – 7.3. ~~Prop~~ Expectation of sums of r.v's

~~Recall. E(X) ... $X = X_1 + \dots + X_n \Rightarrow E[X] = E[X_1] + \dots + E[X_n]$~~

~~If $X_1 =$ ... even if not indep~~

~~Application to counting ...~~

Ex (Deaths in the town { D Bernoulli 1700–1782}) — [Ghahramani Ex. 10.3]

n married couples ~~live~~ are living in a town.
m death occur at random in the town.
Expected # of intact couples?

This ex was already discussed in the lecture covered by Prof. Rudelson

$$X = X_1 + \dots + X_n \quad \text{where} \quad X_i = \begin{cases} 1, & i\text{'th couple is intact} \\ 0, & \text{—} \end{cases}$$

$X_i \sim \text{Bernoulli}(p)$.

$$p = P\{X_i = 1\} = P\{i\text{'th couple is intact}\}$$

$$= P\{\text{all of the } m \text{ death occur among the other } \underbrace{n-1 \text{ couples}}_{= 2n-2 \text{ people}}\}$$

$$= \frac{\binom{2n-2}{m}}{\binom{2n}{m}} = \frac{(2n-2)!}{\cancel{m!}\,(2n-2-m)!} \cdot \frac{\cancel{m!}\,(2n-m)!}{(2n)!}$$

$$= \frac{(2n-m)(2n-m-1)}{(2n)(2n-1)}.$$

$$EX = \sum_1^n E[X_i] = n\,E(X_1) = np = \boxed{\frac{(2n-m)(2n-m-1)}{2(2n-1)}.}$$

Illustration: $n = 1{,}000$.

m	E(X)
100	902
600	490
1200	160
1500	62
1800	10

$\underline{\underline{\text{Ex}}}$ (Coupon Collecting Problem I) (Ex. 3d)

There are n ~~coupons~~ different types of coupons

Each time one obtains a coupon, it is equally likely to be ~~#~~ of ~~any~~ each type.

Compute Expectated ~~#~~ number of the different coupons among N collected. $\| X$

Applications: Clinical trials – collecting info on side effects of a drug.

$Y = n - X$, # of uncollected coupons.

$E[Y] = ?$

$Y = Y_1 + Y_2 + \dots + Y_n$, where $Y_i = \begin{cases} 1, & \text{coupon of } i\text{'th type is not collected} \\ 0, & \text{—} \end{cases}$

~~$Y_i \sim Bernoulli(p)$~~, $E[Y] = \sum_{i=1}^{N} E[Y_i]$. $E[Y_i] = 1 \cdot p + 0(1-p) = p$, where

$p = P\{\text{coupon of } i\text{'th type is not collected}\} = \Big(\underbrace{1 - \frac{1}{n}}_{\text{collected at one trial}}\Big)^N.$

$\Rightarrow E[Y] = \sum_{i}^{N} p = np = n\left(1 - \frac{1}{n}\right)^N.$

$E(X) = \boxed{n - n\left(1 - \frac{1}{n}\right)^N}$

Asympt. Analysis: $n \to \infty$, $N = tn$

$E[Y] \approx n e^{-N/n} = \boxed{n e^{-t}}$.

$E[Y] \leq 1$ for $t \sim \log n \Rightarrow$ for $\underline{N \sim n \log n}$.

Should Expect a complete collection in time $\underline{N \sim n \log n}$.

Let's verify this: