

8

Expectation and variance in multivariate setting

8.1,8.2

Expectation and Variance for sums of r.v.s

Dependence/independence in expected value

Recall, Fact: $\mathbb{E}[X + Y] = \mathbb{E}X + \mathbb{E}Y$ even if X and Y are dependent.

Really?

Ex) Let X, Y be die rolls (maybe dependent die rolls). Then

$$\mathbb{E}[X + Y] = \mathbb{E}X + \mathbb{E}Y = \underline{3.5 + 3.5 = 7}$$

We now verify this fact in two examples.

Independent case

Roll two dice. Let X = first die roll, Y = second die roll.

Calculate $\mathbb{E}[X + Y]$.

$$P_X(k) = \frac{1}{6} \quad k \in \{1, 2, \dots, 6\}$$

Pmf: $p_{X+Y}(n) = \underline{P_X * P_Y(n)}$

$$p_{X+Y}(2) = \underline{\frac{1}{36}}, p_{X+Y}(3) = \underline{\frac{2}{36}}, \dots, p_{X+Y}(6) = \underline{\frac{5}{36}}, p_{X+Y}(7) = \underline{\frac{6}{36}},$$

$$p_{X+Y}(8) = \underline{\frac{5}{36}}, \dots, p_{X+Y}(12) = \underline{\frac{1}{36}}$$

$$\mathbb{E}[X + Y] = \sum_{n=2}^{12} n p_{X+Y}(n) = \sum_{n=2}^{12} n p(n)$$

$$= 2 \cdot p(2) + 12 p(12) \\ + 3 p(3) + 11 p(11) \\ + \vdots$$

$$+ 7 \cdot p(7) \\ = (2+12) \cdot \frac{p(2)+p(12)}{2} =$$

$$+ (3+11) \cdot \frac{p(3)+p(11)}{2}$$

$$+ \vdots \\ + 7 \cdot p(7)$$

$$p(2) = p(12) = \frac{p(2)+p(12)}{2}$$

$$p(3) = p(11) = \frac{p(3)+p(11)}{2}$$

$$7 \left(\underbrace{p(2) + p(12) + p(3) + p(11) + \dots + p(7)}_1 \right)$$

$$= \boxed{7} \checkmark$$

Now, dependent case

X = die roll, $Y = X$.

$$\mathbb{E}[X + Y] = \underline{\mathbb{E}2X = 2\mathbb{E}X = 2 \cdot 3.5 = 7} \quad \checkmark$$

Proof that $E[X + Y] = EX + EY$ for jointly continuous r.v.'s

Jointly cont r.v.'s have a joint pdf, f .

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

$$= \underbrace{\int \int x f(x, y) dy dx}_I + \underbrace{\int \int y f(x, y) dy dx}_{II}$$

$$I = \int_{-\infty}^{\infty} x \underbrace{\int_{-\infty}^{\infty} f(x, y) dy}_{f_X(x)} dx = \int_{-\infty}^{\infty} x f_X(x) dx = EX$$

Similarly, $II = EY$ ✓

Facts: (Expected value, variance, for multiple r.v.s)

1. $\mathbb{E}[X + Y] = \mathbb{E}X + \mathbb{E}Y$

2. $\mathbb{E}[g(X) + f(Y)] = \mathbb{E}g(X) + \mathbb{E}f(Y)$

3. $\mathbb{E}[X \cdot Y] = \mathbb{E}X \cdot \mathbb{E}Y$

4. $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y)$

} Require independence!

- Which ones require independence?
- Generalizes to n r.v.s, e.g., $\mathbb{E}[X_1 + X_2 + \cdots + X_n] = \mathbb{E}X_1 + \cdots + \mathbb{E}X_n$
- Ex) X, Y are independent die rolls. $\mathbb{E}[X \cdot Y] = \underline{\mathbb{E}X \cdot \mathbb{E}Y = (3.5)^2}$

Note: In general, $\mathbb{E}g(X) \neq g(\mathbb{E}X)$!

Ex) You go to the casino with with 1\$. At the roulette wheel, you bet half of your money on black. Repeat this n times. (Each time you have a $\frac{18}{38}$ chance of winning).

1. What is the expected value of your final amount?

$$X = \text{Final amount} = 1 \cdot X_1 \cdot X_2 \cdots X_n$$

$$X_i = \begin{cases} 3/2 & \text{w.p. } \frac{18}{38} \\ 1/2 & \text{w.p. } \frac{20}{38} \end{cases}$$

$$\begin{aligned} \mathbb{E}X &= \mathbb{E}[X_1 \cdot X_2 \cdots X_n] = \mathbb{E}X_1 \cdot \mathbb{E}X_2 \cdots \mathbb{E}X_n \\ &\quad \begin{array}{c} \uparrow \quad \uparrow \quad \uparrow \\ \text{indep} \end{array} \\ &= (\mathbb{E}X_1)^n = \left(\frac{3}{2} \cdot \frac{18}{38} + \frac{1}{2} \cdot \frac{20}{38} \right)^n \\ &= \left(1 - \frac{1}{38} \right)^n = \left(\left(1 - \frac{1}{38} \right)^{38} \right)^{n/38} \approx (e^{-1})^{n/38} = e^{-n/38} \end{aligned}$$

"On average" after 38 plays, you lose most of your money.

Variance and independence

Option a) Let X, Y be different die rolls.

Option b) Let X, Y be the same die roll ($X = Y$).

In which case is $\text{Var}(X + Y)$ larger?

$$a) \text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) = 2 \text{Var}(X)$$

$$b) \text{Var}(X + Y) = \text{Var}(X + X) = \text{Var}(2X) = 4 \text{Var}(X)$$

Fact: Let $X_1, X_2, \dots, X_n$ be independent, identically distributed, r.v.s with $\mathbb{E}X_i = \mu$, $\text{Var}(X_i) = \sigma^2$. Then the average,

$$\bar{X}_n = \frac{X_1 + X_2 + \dots + X_n}{n}$$

satisfies $\mathbb{E}\bar{X}_n = \underline{\mu}$, $\text{Var}(\bar{X}_n) = \underline{\frac{\sigma^2}{n}}$.

Ex) Flip a coin 100 times. Let X be the proportion of heads.

fair

$$\mathbb{E}X = \underline{\frac{1}{2}},$$

$$\text{Var}(X) = \underline{\frac{\text{Var}(\text{Bern}(\frac{1}{2}))}{100}}$$

$$> \boxed{\frac{\frac{1}{2} - \frac{1}{2}^2}{100}}$$

Indicator random variables

Def: Let A be an event. The indicator of that event, 1_A is a r.v. satisfying

$$1_A = \begin{cases} 1, & \text{if } A \text{ occurs} \\ 0, & \text{else.} \end{cases}$$

- $1_A \sim \text{Bern}(p)$, where $p = \underline{P(A)}$.
- $\mathbb{E}1_A = \underline{p = P(A)}$

Ex) Flip a coin 100 times. Let X be the number of runs of 5 heads in a row. ~~What is the distribution of X ?~~ $\mathbb{E} X = ?$

intractable

$$X = \sum_{k=1}^{96} \mathbb{1}_{A_k}$$

$A_k = \{ \text{There is a run of 5 heads starting on flip } k \}$

$$\begin{aligned} \mathbb{E} X &= \sum_{k=1}^{96} \mathbb{E} [\mathbb{1}_{A_k}] = \sum_{k=1}^{96} P(A_k) = \sum_{k=1}^{96} P(\text{HHHHH}) = \sum_{k=1}^{96} 1/32 \\ &= 96/32 = \boxed{3} \end{aligned}$$

Ex) There are 50 people at a party. Each pair are friends with probability $\frac{1}{10}$ (independently of other pairs). On average, how many groups of 4 people where everyone is friends with each other are there?

X

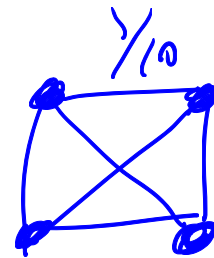
$$X = \sum_{i,j,k,l} \mathbb{1}_{A_{i,j,k,l}}$$

$$\mathbb{E} X = \sum_{\text{groups of 4}} \mathbb{E} [\mathbb{1}_{A_{i,j,k,l}}]$$

$\underbrace{\hspace{10em}}_{P(A_{i,j,k,l}) = \frac{1}{10^6}}$

$$= \boxed{\binom{50}{4} \cdot \frac{1}{10^6}}$$

$$A_{i,j,k,l} = \left\{ \begin{array}{l} \text{People } i,j,k,l \text{ make} \\ \text{a "friend group"} \end{array} \right\}$$



$$\frac{1}{10^6}$$

Ex) What is the expected number of kings in a poker hand?

$$X = \# \text{ of kings} = \mathbb{1}[\text{1 king}] + 2 \cdot \mathbb{1}[\text{2 kings}] + 3 \cdot \mathbb{1}[\text{3 kings}] + 4 \cdot \mathbb{1}[\text{4 kings}]$$

True, but not easiest decomposition to use

$$= \sum_{k=1}^5 \mathbb{1} A_k$$

$$A_k = \{k^{\text{th}} \text{ card is a king}\}$$

$$\mathbb{E} X = \sum_{k=1}^5 \mathbb{E}[\mathbb{1} A_k] = \sum_{k=1}^5 P(A_k) = \frac{4}{52} + \frac{4}{52} + \frac{4}{52} + \frac{4}{52} + \frac{4}{52} = \boxed{\frac{20}{52}}$$

X

Ex) What is the expected number of pairs in a poker hand? (3 of a kind will count for 3 pairs, 4 of a kind for 6 pairs.)

$$X = \sum_{i,j} \mathbb{1}_{A_{ij}}$$

$$A_{i,j} = \{ \text{Cards } i, j \text{ have same rank} \}$$

$$\mathbb{E} X = \sum_{i,j} \mathbb{E} [\mathbb{1}_{A_{ij}}]$$

$$P(A_{ij}) = P(\text{two cards make a pair}) = \frac{3}{51}$$

$$= \boxed{\binom{5}{2} \cdot \frac{3}{51}}$$