Jan 24

2.6. Variance

Def X: rv with mean (= expected value) μ. <u>Variance</u> of X is

$$\mathrm{Var}(X) = E[(X-\mu)^2]$$

<u>Standard deviation</u> of X is

$$\sigma_X = \sqrt{\mathrm{Var}(X)}$$

More conveniently:

$$\begin{aligned}\mathrm{Var}(X) &= E[X^2 - 2\mu X + \mu^2] \\ &= E[X^2] - 2\mu \underbrace{E[X]}_{\mu} + \mu^2 \qquad \text{(by the } \sim\!\sim\text{properties}\sim\!\sim \text{ linearity of expectation)} \\ &= E[X^2] - \mu^2 \qquad \Rightarrow \text{we proved:}\end{aligned}$$

Prop $\boxed{\mathrm{Var}(X) = E[X^2] - (E[X])^2}$

not the same!

Ex (Prediction) Let Y be a r.v.

What is the ~~value~~ constant t that best predicts X? i.e. minimizes the <u>risk</u>

~~To answer this, we minimize the risk,~~

$$R(t) = E[(X-t)^2] \quad ?$$

$$0 = R'(t) = \frac{d}{dt}\left(E[X^2] - 2t\,E[X] + t^2\right) = -2E[X] + 2t$$

$$\Rightarrow \quad t = E[X]$$

Thus. <u>the value that best predicts X is $E[X]$</u>

(and the associated risk is <u>$\mathrm{Var}(X)$</u>).

Prop (properties of var)

$$\boxed{\begin{aligned}&(i)\ Var(aX) = a^2\, Var(X), \quad \text{so} \quad \sigma_{aX} = a\cdot\sigma_X \\ &(ii)\ Var(X+b) = Var(X), \quad \text{so} \quad \sigma_{X+b} = \sigma_X\end{aligned}}$$

Proof of (ii)

$$Var(X+b) = E\left[\left((X+b) - E(X+b)\right)^2\right] = E\left[(X+b-E[X]-b)^2\right] \quad \text{(by prop's of expectation p.33)}$$

$$= E\left[(X-E[X])^2\right] = Var(X) \qquad \text{QED.}$$

Warning It is NOT true in general that $Var(X+Y) = Var(X)+Var(Y)$ (e.g. for $Y=-X$).

Ex (The matching problem)

N homeworks are returned to N students at random; $X = \#$ students receiving their own HW

~~How many students will get their own homework on average?~~

~~Recall that we calculated before P{at least 1 student gets his/her own HW} ≈ 1 − 1/e, see p. 17~~

We computed before that $\mathbb{P}(X\geq 1)\simeq 1-1/e$; $\mathbb{E}[X]=1$. $\boxed{Var(X)=?}$

~~Method of indicators~~

Represent $X = \sum_{i=1}^{N} X_i$ where $X_i = \begin{cases} 1, & \text{if student } i \text{ gets his/her own HW} \\ 0, & \text{otherwise} \end{cases}$

"indicators"

~~$E[X] = \sum_{i=1}^N E[X_i]$~~

~~$E[X_i] = 1\cdot P\{X_i=1\} + 0\cdot P\{X_i=0\} = P\{X_i=1\} = P\{\text{student } i \text{ gets own HW}\} = \frac{1}{N}$ (why?)~~

~~Therefore $E[X] = N\cdot\frac{1}{N} = 1$~~

~~Ex What is the standard deviation of the students who get their own hw?~~

$$E[X^2] = E\left[\left(\sum_{i=1}^N X_i\right)^2\right] = E\left[\sum_{i=1}^N X_i^2 + \sum_{i\neq j} X_iX_j\right] = \sum_{i=1}^N E[X_i^2] + \sum_{i\neq j} E[X_iX_j].$$

$E[X_i^2] = E[X_i] = \frac{1}{N}$, as before

$E[X_iX_j] = 1\cdot P\{X_iX_j=1\} + 0\cdot P\{X_iX_j=0\} = P\{X_i=1 \text{ and } X_j=1\} = P\{\text{both students } i,j \text{ get their own HW}\}$

$= \frac{1}{N(N-1)}$ (Why? e.g. by conditioning on one of students)

Therefore

$$E[X^2] = N\cdot\frac{1}{N} + \underbrace{N(N-1)}_{\#\text{ pairs of } i\neq j}\cdot\frac{1}{N(N-1)} = 2$$

Hence $Var(X) = E[X^2] - (E[X])^2 = 2-1 = 1$ $\qquad \sigma_X = \sqrt{1} = 1$

2.4.

Prop (independence) Let X, Y be indep. discrete r.v.'s. Then

(i) $\mathbb{E}[XY] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$;

(ii) $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$

(i) ~~Let PMF be $\mathbb{P}\{X=x_i\}=p_i$, $\mathbb{P}\{Y=y_j\}=q_j$~~
Let X take values $\{x_i\}$, Y take values $\{y_j\}$. Then

$$\mathbb{E}[XY] = \sum_{i,j} (x_i y_j)\, \mathbb{P}\{X=x_i, Y=y_j\} \quad \text{(by Remark)}$$

$$= \sum_{i,j} x_i y_j\, \mathbb{P}\{X=x_i\}\cdot\mathbb{P}\{Y=y_j\} \quad \text{(by independence)}$$

$$= \Big(\sum_i x_i\, \mathbb{P}\{X=x_i\}\Big)\Big(\sum_j y_j\, \mathbb{P}\{Y=y_j\}\Big) \quad \text{(by absolute convergence)}$$

$$= \mathbb{E}[X]\cdot\mathbb{E}[Y] \quad \text{(by def.)}$$

(ii)
$$\mathrm{Var}(X+Y) = \mathbb{E}[(X+Y)^2] - (\mathbb{E}[X+Y])^2$$

$$= \mathbb{E}[X^2 + 2XY + Y^2] - (\mathbb{E}[X]+\mathbb{E}[Y])^2$$

$$= \mathbb{E}[X^2] + 2\mathbb{E}[X]\cdot\mathbb{E}[Y] + \mathbb{E}[Y^2] - \big(\mathbb{E}[X]^2 + 2\mathbb{E}[X]\cdot\mathbb{E}[Y] + \mathbb{E}[Y]^2\big) \quad \text{(indep.)}$$

$$= \big(\mathbb{E}[X^2] - \mathbb{E}[X]^2\big) + \big(\mathbb{E}[Y^2] - \mathbb{E}[Y]^2\big)$$

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

~~Ex Alice & Bob want to cut out a re~~
~~Alice chooses a number X at random & cuts a square piece of paper with sides X.~~
~~Bob chooses two numbers Y, Z independently, and with the same distr. as X and cuts a rectangular piece of paper with~~

2.8. Bernoulli and Binomial distributions

Bernoulli

- Experiment has two outcomes: success, failure. $P\{\text{success}\} = p$, $P\{\text{failure}\} = 1-p$.

Let r.v. X be the indicator of success, then X takes two values: $\begin{cases} 1 \text{ with prob } p, \\ 0 \text{ with prob. } 1-p \end{cases}$

Formally:

Def A Bernoulli r.v with parameter p is a r.v X with pmf

$$P\{X=1\} = p, \quad \cancel{p(0) = 1-p} \quad P\{X=0\} = 1-p.$$

Notation: $X \sim \text{Bernoulli}(p)$.

- $E[X] = 1\,p + 0\cdot(1-p) = \underline{p}$
- $E[X^2] = E[X] = p;$ $\quad Var(X) = p - p^2 = \underline{p(1-p)}$

Binomial

- Experiment consists of n independent trials, each having success with prob p, failure with prob. $1-p$.

Let r.v X be the # of successes

- pmf: $\quad P\{X=k\} = P\{k \text{ successes in } n \text{ trials}\} = \underbrace{\binom{n}{k}}_{\substack{\#\text{ of possible arrangements} \\ \text{of successes \& failures}}} \underbrace{p^k (1-p)^{n-k}}_{\substack{\text{Prob of each} \\ \text{arrangement.}}}$

SSFFSFFFSSFSF

Formally,

Def A Binomial r.v with parameters n, p is a r.v. X with pmf

$$P\{X=k\} = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \ldots, n$$

Notation:

$$X \sim \text{Binom}(n, p)$$

Prop Let $X \sim \text{Binom}(n,p)$. Then
$$\mathbb{E}[X] = np, \qquad \text{Var}(X) = np(1-p).$$

$X = \sum_1^n X_i$ where $X_i \sim \text{Bernoulli}(p)$ indep. (indicators of successes).

$\Rightarrow \quad \mathbb{E}[X] = \sum_1^n \underbrace{\mathbb{E}[X_i]}_{= p} = np.$

$\text{Var}(X) = \sum_1^n \text{Var}(X_i)$ (independence)

$= n \cdot p(1-p).$

Ex In a class of 40 students, on average 2 students are sick. What is the prob. that 4 students are sick?

$X = \#$ sick students $\sim \text{Binom}(40, p)$

$\mathbb{E}[X] = 40 \cdot p = 2 \quad \Rightarrow \quad p = \frac{1}{20}.$

$P\{X=4\} = \binom{40}{4}\left(\frac{1}{20}\right)^4\left(1-\frac{1}{20}\right)^{36} \approx 0.0901.$

Ex A fair coin is tossed n times.

$X = \#\text{heads} \sim \text{Binom}(n, \frac{1}{2})$

$\Rightarrow \mathbb{E}[X] = \frac{n}{2}, \quad \text{Var}(X) = \frac{n}{4} \quad \Rightarrow \quad \sigma_X = \frac{\sqrt{n}}{2}.$

Thus, typically there are $\frac{n \pm \sqrt{n}}{2}$ heads

E.g. for $n = 10^6$ tosses, $500{,}000 \pm 500$ heads.

2.8 (contd.) Poisson distribution

- Recall: $Y \sim \text{Binom}(n,p)$ if $Y = \#$ successes in n indep trials, where $p =$ prob. of success in each trial.

  pmf: $p(k) = \binom{n}{k} p^k (1-p)^{n-k}$, $k = 0, 1, \dots, n$

  $E[Y] = np$, $\text{Var}(Y) = np(1-p)$.

Ex In a class of 40 students, on average 2 students are sick. What is the prob that 4 students are sick?

$X = \#$ sick students $\sim \text{Binom}(\underset{n}{40}, p)$. $E[X] = 40 \cdot p = 2 \Rightarrow p = 1/20$

$p(4) = \binom{40}{4}\left(\frac{1}{20}\right)^4\left(1-\frac{1}{20}\right)^{36} = 0.09012$

2.8 (contd.) Poisson Approximation to Binomial

- The pmf of Binom(n,p) is sometimes difficult to compute, especially for large n.

- $\Rightarrow$ approximate. Assume that successes are rare (like in the last example):

$$\underbrace{n \to \infty}_{\text{many trials}}, \quad \underbrace{np \to \lambda = \text{const}}_{\text{const \# of successes on ave}} \qquad (\text{thus } p \to 0)$$

- Let us approximately compute $P\{X=k\}$ $\left(\underset{p(n)}{p} = \lambda/n\right)$

$$p(k) = \frac{n!}{k!(n-k)!}\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k}$$

$$= \frac{n(n-1)\cdots(n-k+1)}{k!}\,\frac{\lambda^k}{n^k}\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-k}.$$

Note: (1) $\frac{n(n-1)\cdots(n-k+1)}{n^k} \to 1$ $(n \to \infty,\ k \text{ fixed})$

(2) $\left(1-\frac{\lambda}{n}\right)^n \to e^{-\lambda}$ $(n \to \infty)$ $\left[\text{Recall } \lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n = e^x \text{ for each } x \in \mathbb{R}\right]$

(3) $\left(1-\frac{\lambda}{n}\right)^{-k} \to 1$ $(n \to \infty,\ k \text{ fixed})$

Hence

$$\boxed{p(k) \approx \frac{\lambda^k}{k!}e^{-\lambda}}$$

Def (Poisson distr.) A r.v. X has Poisson distribution with parameter $\lambda > 0$ if X has ~~mpf~~ pmf

$$\mathbb{P}\{X=k\} = e^{-\lambda} \cdot \frac{\lambda^k}{k!}, \quad k=0,1,2,\dots$$

Denoted $X \sim \text{Poisson}(\lambda)$.

We have proved:

THM (Poisson Limit Thm) Let $X_n \sim \text{Binom}(n, \frac{\lambda}{n})$.

Then, as $n \to \infty$, the pmf of X_n converges pointwise to the pmf of Poisson(λ), i.e.

$$\mathbb{P}\{X_n = k\} \to e^{-\lambda} \frac{\lambda^k}{k!} \qquad \forall\, k=0,1,2,\dots$$

Remark Since $\sum \mathbb{P}\{X=k\}=1$, we gave a probabilistic proof of the identity $\boxed{\sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{\lambda}}$

Cor If $X \sim \text{Poisson}(\lambda)$ then

$$\mathbb{E}[X] = \lambda, \qquad \text{Var}(X) = \lambda.$$

Informal proof: $X \approx Y \sim \text{Binom}(n, \frac{\lambda}{n})$ for large λ.

$$\Rightarrow \mathbb{E}[X] \approx \mathbb{E}[Y] = n \cdot \frac{\lambda}{n} = \lambda$$

$$\text{Var}(X) \approx \text{Var}(Y) = n \frac{\lambda}{n}\left(1 - \frac{\lambda}{n}\right) \to \lambda \text{ as } n \to \infty.$$

Ex Make this argument formal. Use the following tool:

Dominated Convergence Thm for series: Let $a_{n,k}$ be numbers; $n,k \in \mathbb{N}$

Suppose:
- $\lim_{n\to\infty} a_{n,k} = a_k \quad \forall k$
- $|a_{n,k}| \le b_k \quad \forall n$ and $\sum_{k=1}^{\infty} b_k < \infty$.

Then $\lim_{n\to\infty} \sum_{k=1}^{\infty} a_{n,k} = \sum_{k=1}^{\infty} \lim_{n\to\infty} a_{n,k} = \sum_k a_k$ (the series converges)

Remark Use Poisson distr. for rare successes, e.g.

\# typos on a page of a book

\# software failures on a given day, ...

Ex (in Ex. p.34): of 40 students, ~~...~~ 2 are sick on ave. $X \doteq$ # sick.

- $X \approx \text{Poisson}(2) \Rightarrow P\{X=4\} \approx e^{-2} \cdot \frac{2^4}{4!} = 0.0902$ (compare to 0.0901)

~~Ex~~

- $P\{\text{no sick students}\} = P\{X=0\} = e^{-2} = 0.14$.
- $P\{\text{at least } \cancel{\text{one}}\ \text{two sick}\} = 1 - P\{X=0\} - P\{X=1\} = 1 - e^{-2} - 2e^{-2} = 0.56$.