

MATH-235 Probability and Statistics
With Mats Stensrud

Lecture 1

Plan:

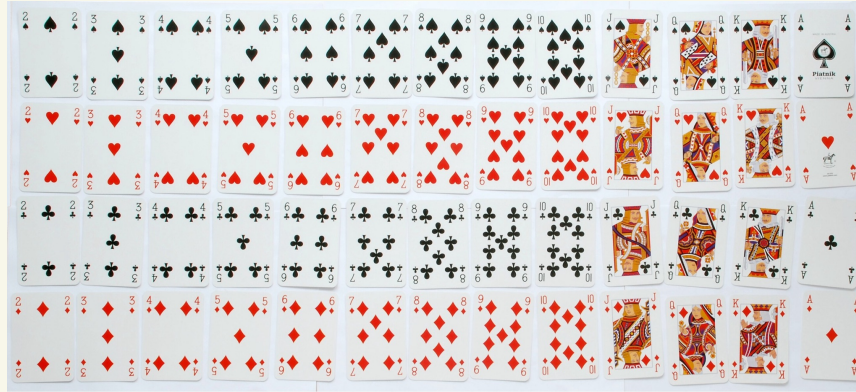
- Intro / Info
- Naïve definition of probability
- Counting
- Story proofs



Fermat 1600's

Probability

• Built on sets



Experiment is performed.

- A sample space is the set of all possible outcomes of an experiment. (S)
- An event is a subset of the sample space.
(A or B).

- S is sample space, then the events $A, B \subseteq S$.
- $A \cup B$ is the event that A or B occurred
- $A \cap B$ is the event that both A and B occur.
- Complement A^c is the event that occurs if and only if A does not occur.
- $(A \cup B)^c = A^c \cap B^c$ and $(A \cap B)^c = A^c \cup B^c$ (De Morgan's law).

Example cards

- Pick a card
- S is the set of 52 cards
- Consider the events:

A : card is an ace

B : black suit

D : Diamond

H : Heart

$A \cap H$: Ace of Heart

$A \cap B$: Ace of Spades, Ace of Clubs

$A \cup D \cup H$: red card or ace

$(A \cup B)^c$: red non-ace

of events 2^{52}

Naive def of probability

$$P(A) = \frac{\# \text{ of favourable outcomes (in } A)}{\# \text{ of possible outcomes (in } S)} = \frac{|A|}{|S|}$$

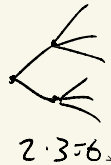
$|A|$ is the size of A .

Multiplication rule for counting

r experiments.

Each has n_r outcomes.

overall there are $n_1 n_2 \dots n_r$ possible outcomes



Example Poker

Binomial coefficient: $\binom{n}{k} = \frac{n(n-1) \cdots (n-k+1)}{k!} = \frac{n!}{(n-k)!k!}$ or 0 if $k > n$

Full house

3 cards of one rank, 2 cards of another rank.

$$\frac{13 \binom{4}{3} \cdot 12 \binom{4}{2}}{\binom{52}{5}} = \frac{|A|}{|S|}$$

	order matters	order doesn't matter
Replace	n^k	$\binom{n+k-1}{k}$
Don't replace	$n(n-1) \cdots (n-k+1)$	$\binom{n}{k}$

Ex 10 people. Divide into teams of 6 and 4.

$$\binom{10}{4} = \binom{10}{6}$$

Choose two teams of 5.

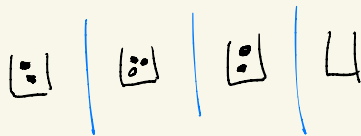
$$\binom{10}{5} / 2.$$

Ex pick k times from n objects; order doesn't matter, with replacement

$\binom{n+k-1}{k}$ ways

$$\begin{aligned} n &= 4 \\ k &= 7 \end{aligned}$$

n boxes



positions $n+k-1$

Q: How many ways can we "organise" the separators (or balls)

$$\binom{n+k-1}{k} = \binom{n+k-1}{n-1}$$

$$k=0: \Rightarrow \binom{n-1}{0} = 1.$$

$$k=1: \Rightarrow \binom{n}{1} = n$$

$$n=2 \Rightarrow \binom{k+1}{k} = \binom{k+1}{1} = k+1.$$

Story proof

- Proof by interpretation
- Eg. for counting, count in different ways
- Benefit: don't need to do tedious algebra.

Ex choosing the complement

For non-neg. integers $k \leq n$, $\binom{n}{k} = \binom{n}{n-k}$

Ex (Captain of team)

n people

k team members

1 captain

$$\binom{n}{k} k = n \binom{n-1}{k-1}$$

Ex Vandermonde's identity

m : first year students

n : second year students

k : # in a committee

$$\binom{m+n}{k} = \sum_{j=0}^k \binom{m}{j} \binom{n}{k-j}$$

Non-naive Definition of probability

We need two objects:

S , the sample space.

P , a function whose argument is an event $A \subseteq S$,

returns $P(A) \in [0,1]$ as output.

Two axioms:

(1) $P(\emptyset) = 0$, $P(S) = 1$

(2) $P\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} P(A_n)$ if $A_1, A_2, \dots$ are disjoint (mutually exclusive)

Lecture 2

- non-naive definition of probability
- Inclusion-exclusion
- Independence
- Conditional prob.

Ex Birthday problem.

$k > 365$, then prob. 1.

$\dot{\square} \dot{\square} \ddot{\square} \dot{\square} \dots$

$k \leq 365$

$$P(\text{no match}) = \frac{365 \cdot 364 \cdot 363 \cdot \dots \cdot (365 - k + 1)}{365^k}$$

$$P(\text{match}) = 1 - P(\text{no match})$$

$$P(\text{match}) = \begin{cases} 50.7 & \text{if } k=23 \\ \approx 99.9 & \text{if } k=100. \end{cases}$$

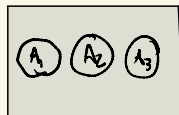
$$\binom{k}{2} = \frac{k(k-1)}{2}$$

$$k=23 \Rightarrow \frac{23 \cdot 22}{2} = 253.$$

Axioms p is the probability, a non-negative real number,

(1) $P(\emptyset) = 0$, $P(S) = 1$.

(2) $P(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} P(A_n)$ if $A_1, A_2, \dots$ are disjoint

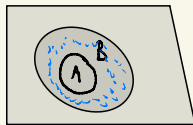


Properties

(1) $P(A^c) = 1 - P(A)$ Proof: $1 = P(S) = P(A \cup A^c) = P(A) + P(A^c)$ disjoint: $A \cap A^c = \emptyset$

(2) A is contained in B , $A \subseteq B$

$$P(A) \leq P(B)$$

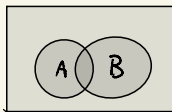


Proof: $B = A \cup (B \cap A^c)$

$$P(B) = P(A) + P(B \cap A^c) \quad \text{by axiom 2.}$$

$$\geq P(A)$$

(3) If A and B are not disjoint



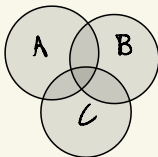
$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

Proof: $P(A \cup B) = P(A \cup (B \cap A^c)) = P(A) + P(B \cap A^c) \stackrel{*}{=} P(A) + P(B) - P(A \cap B)$

Note that

$$(*) P(B) = P(B \cap A) + P(B \cap A^c)$$

More generally, Exclusion-inclusion:



$$\begin{aligned}
 P(A \cup B \cup C) &= P(A) + P(B) + P(C) \\
 &\quad - P(A \cap B) - P(A \cap C) - P(B \cap C) \\
 &\quad + P(A \cap B \cap C)
 \end{aligned}$$

Thm (inclusion-exclusion)

$$\begin{aligned}
 &P(A_1 \cup A_2 \cup \dots \cup A_n) \\
 &= \sum_{i=1}^n P(A_i) - \sum_{1 \leq i < j \leq n} P(A_i \cap A_j) + \sum_{1 \leq i < j < k \leq n} P(A_i \cap A_j \cap A_k) \\
 &\quad - \dots + (-1)^{n+1} P(A_1 \cap A_2 \cap \dots \cap A_n).
 \end{aligned}$$

Example Montmort's problem (Matching problem)

n cards, labelled $1, 2, \dots, n$.

A_j is the event that j 'th card matches.

our probability of interest: $P(A_1 \cup A_2 \cup \dots \cup A_n)$.

$$P(A_j) = \frac{1}{n}.$$
$$P(A_1 \cap A_2) = \frac{(n-2)!}{n!} = \frac{1}{n(n-1)}.$$
$$P(A_j \cap A_i) = \frac{(n-2)(n-3) \dots 1}{n \cdot (n-1)(n-2)(n-3) \dots 1}$$

$$P(A_1 \cap A_2 \cap \dots \cap A_k) = \frac{(n-k)!}{n!} = \frac{1}{n(n-1)(n-2) \dots (n-k+1)}$$

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = \sum_{j=1}^n P(A_j) - \sum_{1 \leq i < j \leq n} P(A_i \cap A_j) + \dots$$
$$= 1 - \frac{n(n-1)}{2} \cdot \frac{1}{n(n-1)} + \frac{n(n-1)(n-2)}{3!} \cdot \frac{1}{n(n-1)(n-2)} - \dots = 1 - \frac{1}{2!} + \frac{1}{3!} - \frac{1}{4!} + \dots + (-1)^{n+1} \frac{1}{n!}$$
$$\approx 1 - \frac{1}{e}.$$

Independence

- Intuition: we want to say that one event does not give information about another event.

Def. A and B are independent if

$$P(A, B) := P(A \cap B) = P(A) \cdot P(B)$$

A, B, C

$$P(A \cap B) = P(A)P(B)$$

$$P(A \cap C) = P(A)P(C)$$

$$P(C \cap B) = P(B)P(C)$$

$$P(A \cap B \cap C) = P(A)P(B)P(C)$$

Generally, $A_1, A_2, \dots, A_n$ are independent if for every subset of

these events
$$P\left(\bigcap_{i \in I} A_i\right) = \prod_{i \in I} P(A_i)$$

Ex: Newton-Pepys problem.

What is most likely:

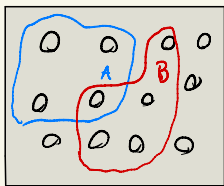
- (A) At least one six with six dice
- (B) At least two six's with 12 dice
- (C) At least three six's with 18 dice

$$P(A) = 1 - P(A^c) = 1 - \left(\frac{5}{6}\right)^6 = 0.665$$

$$P(B) = 1 - \binom{12}{0} \left(\frac{5}{6}\right)^{12} - \binom{12}{1} \left(\frac{1}{6}\right) \left(\frac{5}{6}\right)^{11} = 0.619.$$

$$P(C) = 1 - \binom{18}{0} \left(\frac{5}{6}\right)^{18} - \binom{18}{1} \left(\frac{1}{6}\right) \left(\frac{5}{6}\right)^{17} - \binom{18}{2} \left(\frac{1}{6}\right)^2 \left(\frac{5}{6}\right)^{16} = 0.597$$

Conditional probability



Def: $P(A|B) = \frac{P(A \cap B)}{P(B)}$ if $P(B) > 0$.

Intuition: what happens if I do an experiment infinitely many times?

Ex: Pick stones from B.

check how many times A occurred.

Compute the fraction

Thm: $P(A, B) = P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)$

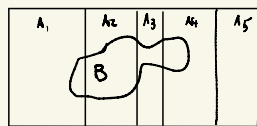
$P(A_1 \cap A_2 \dots \cap A_n) = P(A_1)P(A_2|A_1) \dots P(A_n|A_1 \cap A_2 \dots \cap A_{n-1})$

Bayes' thm

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}$$

Partitioning is useful!

Create $A_1, A_2, \dots, A_n$ to be disjoint.



$$\begin{aligned} P(B) &= P(B \cap A_1) + P(B \cap A_2) + \dots \\ &\quad + P(B \cap A_n) \\ &= P(B|A_1)P(A_1) + \dots \\ &\quad P(B|A_n)P(A_n) \end{aligned}$$

"Law of total probability"

Lecture 3

Plan

- Conditional probability ("|")
- Conditional independence
- Examples:
 - Monty Hall
 - Simpson's

Example

Get two cards from std. deck

(1) $P(\text{both aces} \mid \text{have one ace})$

(2) $P(\text{both aces} \mid \text{ace of spades})$

Case 1

$$\frac{P(\text{both aces} \cap \text{have one ace})}{P(\text{have one ace})} = \frac{\frac{\binom{4}{2}}{\binom{52}{2}}}{1 - \frac{\binom{48}{2}}{\binom{52}{2}}} = \frac{1}{33}$$

Case 2



$$\frac{3}{51} \approx \frac{1}{17}$$

Example (Disease testing)

- Disease affects 1% of the pop.
- Test positive means "disease detected".
- Company says 95% accurate.
- D is event that a patient has a disease
- T patient tests positive.

Suppose: $P(T|D) = 0.95 = P(T^c|D^c)$

Patient's interest is: $P(D|T)$.

$$P(D|T) = \frac{P(T|D)P(D)}{P(T)} = \frac{P(T|D)P(D)}{P(T|D)P(D) + P(T|D^c)P(D^c)} = \frac{0.95 \cdot 0.01}{0.95 \cdot 0.01 + 0.05 \cdot 0.99} = \underline{\underline{0.16}}$$

Caution!

* Distinguish $P(A|B)$ and $P(B|A)$. Prosecutor's fallacy.

Ex: Sally Clark.

Suppose SIDS affects $\frac{1}{8500}$

Naive calculation: $\frac{1}{8500} \cdot \frac{1}{8500} \approx \frac{1}{70 \cdot 10^6}$

- Fetal Alcohol syndrome (FAS)

- A: Event having FAS

- B: Drinking in first trimester

$P(B|A)$.

* Confuse $P(A)$ and $P(A|B)$. $P(A|A)=1$.
"prior" "updated", "posterior"

* Conditional independence vs independence

Def Conditional independence.

Events A, B are conditionally independent given C if

$$P(A, B | C) = P(A | C) P(B | C).$$



$$P(A, B) = P(A)P(B)$$

Example Chess /

A : Event I win first game
 B : Event I win second game

C : Event that Geller is strong

$$P(A \cap B) = P(A)P(B|A)$$

$$\text{but } P(B|A) \neq P(B)$$

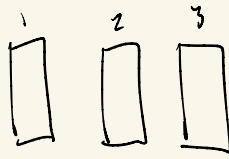
Example Genetics /

- G event that a person has KRAS.
- S event that a person smokes
- C event that a person has Cancer.

$$P(G, S) = P(G)P(S)$$

$$P(G, S | C) \neq P(G | C) P(S | C)$$

Ex Monty Hall



3 doors.

2 have goats behind them

1 have a car.

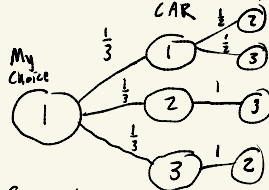
Game:

- You choose a door

- Monty Hall opens another door, shows a goat.

- Should we switch?

~ Suppose I chose door 1.



S: Event winning car when I switch.
 D_j : Event car is behind door j.

$$P(S) = \underbrace{P(S|D_1)}_{=0} \overset{P(D_1)}{\overset{1}{\parallel}} \frac{1}{3} + \underbrace{P(S|D_2)}_{=1} \frac{1}{3} + \underbrace{P(S|D_3)}_{=1} \frac{1}{3}$$
$$= \frac{2}{3}.$$

Ex Simpson's paradox

	Beatrice		David	
	Heart transplant	Remove nail	Heart transplant	Remove nail
Success	70	10	2	81
Failures	20	0	8	9

A: Successful surgery

B: Beatrice did the surgery

C: Heart transplant.

$$P(A|B) = \frac{80}{100} = 0.8 < P(A|B^c) = 0.83$$

$$P(A|B, C) = \frac{7}{9}$$

$$> P(A|B^c, C) = \frac{2}{10}$$

$$P(A|B, C^c) = 1$$

$$> P(A|B^c, C^c) = \frac{81}{90}$$

$$P(A|B) = P(A|C, B) \cdot P(C|B) + P(A|C^c, B) \cdot P(C^c|B)$$

$$P(A|B^c) = P(A|C, B^c) \cdot P(C|B^c) + P(A|C^c, B^c) \cdot P(C^c|B^c)$$

Lecture 4

Plan

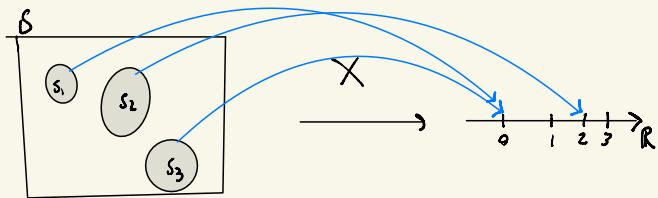
- Random variables
- PMF, CDF
- Bernoulli, Binomial, Hypergeometric

What is a variable?

$$X + Z = 9$$

$$X = 7$$

$R.V. (Random variable) X$: function from the sample space S to the real line.



Ex: Fair coin, flipped two times.

$$S = \{HH, HT, TH, TT\}$$

Let X be # heads, can take values $\{0, 1, 2\}$.

$$X(HH) = 2, X(HT) = 1, X(TH) = 1, X(TT) = 0$$

$$P(\underbrace{X=2}) = 0.25$$

$$\{X=2\}$$

Event $X=2$

$$P(\underbrace{X=1}) = 0.5$$

$$\{HT, TH\}$$

Def Probability mass function (PMF) of random variable.

Consider a discrete RV X with possible values

$$a_1, a_2, \dots, a_m \quad \text{or} \quad a_1, a_2, \dots$$

The PMF is just $P(X = a_j) = p_j$ for all j and we need that

$$(I) \quad p_j \geq 0$$

$$(II) \quad \sum_j p_j = 1.$$

Def Cumulative Distribution Function (CDF).

Consider RV X , $\underbrace{X \leq x}_{\text{Event}}.$

Define $F(x) = P(X \leq x)$

Ex: two coins, $P(X \leq 1) = 0.75.$

Def. Bernoulli distribution

A RV X has a Bern. distribution if X has only two possible values 0 and 1, and if $P(X=1)=p$ and $P(X=0)=1-p$, where $1 \geq p \geq 0$

Def. Binomial (n, p) distribution

The distribution of # successes in n independent Bern(p) trials, called Binom(n, p), which has distribution (PMF):

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

for $k \in \{0, 1, \dots, n\}$

$n=8$ $\underbrace{1111}_{k=5} 000$

← This is a PMF because:

$$\sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} \stackrel{\text{Binomial Thm}}{=} 1.$$

Interpretation of Binomial

- # successes in n Bern trials
- sum of indicator RVs (indicators of success),

$$X = X_1 + X_2 + \dots + X_n \quad \text{where } X_j = \begin{cases} 1, & \text{if } j\text{th trial is succ.} \\ 0, & \text{otherwise} \end{cases}$$

$X_1, X_2, \dots, X_n$ iid (independent and identically distributed).

Remark Binomial thm.

$$(x+y)^n = \sum_{k=0}^n \binom{n}{k} x^k y^{n-k}$$

$$= (x+y)(x+y)\cdots(x+y)$$

$$(a+b)(c+d) = ac+ad+bc+bd$$

Ex (sum of Binomial RV)

$$X \sim \text{Binom}(n, p) \quad Y \sim \text{Binom}(m, p) \quad X, Y \text{ independent,}$$

↑
distributed as

$$\text{then } X + Y \sim \text{Binom}(n+m, p)$$

Alternative ways of seeing this:

$$(*) \quad X = \sum_{i=1}^n x_i, \quad Y = \sum_{j=1}^m y_j \quad \Rightarrow \quad X + Y = \sum_{i=1}^n x_i + \sum_{j=1}^m y_j.$$

$$\begin{aligned} (* *) \quad P(X + Y = k) &= \sum_{j=0}^k P(X + Y = k \mid X = j) P(X = j) \\ &= \sum_{j=0}^k P(Y = k - j \mid X = j) \binom{n}{j} p^j (1-p)^{n-j} \\ &= \sum_{j=0}^k P(Y = k - j) \binom{n}{j} p^j (1-p)^{n-j} \dots \dots \dots (1 \text{ don't finish}) \end{aligned}$$

Ex 5 cards, find distribution of # aces.

X is RV denoting # Aces.

(zero unless $k=0,1,2,3,4$.)

$$P(X=k) = \frac{\binom{4}{k} \binom{48}{5-k}}{\binom{52}{5}}, \quad k \in \{0,1,2,3,4\}$$

Ex balls in a basket.

w white balls

b black balls

Draw n balls without replacement

X is # white balls.



Then X has a hypergeometric distribution $X \sim \text{HGeom}(w, b, n)$

$$P(X=k) = \frac{\binom{w}{k} \binom{b}{n-k}}{\binom{w+b}{n}}, \quad 0 \leq k \leq w, \quad 0 \leq n-k \leq b. \quad \text{Valid PMF?}$$

$$\sum_{k=0}^n \frac{\binom{w}{k} \binom{b}{n-k}}{\binom{w+b}{n}} = 1$$

by Vandermonde.

Def. Independence of RV.

X and Y are independent if

$$P(X \leq x, Y \leq y) = P(X \leq x) P(Y \leq y) \quad \text{for all } x, y$$

In discrete case:

$$P(X=x, Y=y) = P(X=x) P(Y=y) \quad \text{for all } x, y.$$

Ex. Toss 2 dice.

X is # 1st die

Y is # 2nd die

$(X+Y)$ independent of $(X-Y)$

$$0 = P(X+Y=12, X-Y=1) \neq P(X+Y=12) P(X-Y=1) = \frac{1}{36} \cdot \frac{5}{36}.$$

Lecture 5

Averages

Geometric distribution

Negative Binomial

Examples

Repetition iid.

Ex 1: X result of die 1 $X \in \{1, 2, 3, 4, 5, 6\}$ iid
 Y result of die 2 $Y \in \{1, 2, 3, 4, 5, 6\}$

Ex 2: X same as before

Y : indicator of Geller eating meat
Indep., not iid. distr.

Ex 3: Election, Trump. Dependent and not iid. distributed.

Ex 4: one die, toss n times.

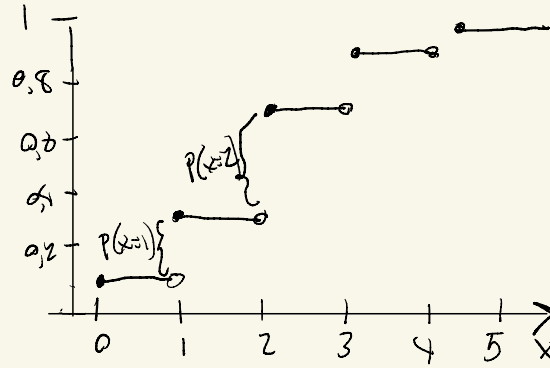
X : # heads, Y : # tails

Dependent and identically distributed.

CDF.

$$P(X \leq x) = F(x)$$

$$\text{Ex: } X \sim \text{Binom}(n=4, p=0.5)$$



$$P(1 < X \leq 3) = P(X \leq 3) - P(X \leq 1) = F(3) - F(1)$$

$$P(X \leq 1) + P(1 < X \leq 3) = P(X \leq 3)$$

CDF is (1) non-decreasing, (2) right continuous (3) $F(x) \rightarrow 0$ when $x \rightarrow -\infty$
 $F(x) \rightarrow 1$ when $x \rightarrow \infty$

Expectation

$$\frac{1 + 2 + 3 + 4 + 5 + 6}{6} = 3.5$$

Gauss:

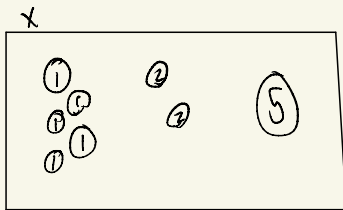
$$\sum_{i=1}^n i = (n+1) \frac{n}{2}$$

Def Expectation / Expected value / mean / average

$$\mathbb{E}(X) = \sum_x x \underbrace{P(X=x)}_{\text{PMF}}, \text{ for those } x \text{ with } P(X=x) > 0.$$

Notation: $\mathbb{E}X = \mathbb{E}(X)$

$$\mathbb{E}X^2 = \mathbb{E}(X^2) = \cancel{\mathbb{E}(X)^2}$$



$$E(x) = \frac{5}{8} 1 + \frac{2}{8} 2 + \frac{1}{8} 5$$

$$= \frac{1}{8} (1+1+1+1+2+2+5)$$

Thm linearity of Exp.

X, Y are RV, defined on the same sample space.

$$E(X+Y) = E(X) + E(Y), \quad cE(X) = E(cX).$$

$$\begin{aligned} & \sum_x x P(X=x) + \sum_y y P(Y=y) \\ &= \sum_s X(s) P(\{s\}) + \sum_s Y(s) P(\{s\}) \\ &= \sum_s (X+Y)(s) P(\{s\}) = E(X+Y) \end{aligned}$$

Ex. $X \sim \text{Bern}(p)$

$$E(X) = 1 P(X=1) + 0 P(X=0) = p$$

Let X indicate that the event A occurs:

$$X = \begin{cases} 1 & \text{if } A \text{ occurs} \\ 0 & \text{otherwise} \end{cases}$$

$$P(A) = E(X) \quad \text{"Fundamental bridge"}$$

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

Difficult

$$\mathbb{E}(X) = \sum_{k=0}^n k P(X=k) = \sum_{k=0}^n k \binom{n}{k} p^k q^{n-k}$$

$\vdots$

$1-p=q$.

Easy way

$$I_i \sim \text{Bern}(p) \quad i \in \{1, 2, \dots, n\}$$

$$X = I_1 + I_2 + \dots + I_n$$

$$\mathbb{E}(I_i) = p$$

$$\mathbb{E}(X) = np$$

Ex. hypergeometric distribution

5 cards, X_j be indicator RV of j 'th being an ace.

$$E(X_5) = E(X_1) \quad \text{symmetry.}$$

$$E(X_1 + X_2 + \dots + X_5) = E(X_1) + E(X_2) + \dots + E(X_5) = 5 \cdot \frac{4}{52} = \frac{5}{13}.$$

Geometric distribution (p. 157-158)

Sequence of $\text{Bern}(p)$ trials, stop at first success.

$X = \#$ of failures before 1st success. $0 \leq q = 1-p \leq 1$

$$P(X=k) = q^k p$$

Valid PMF?

$$\sum_{k=0}^{\infty} q^k p = p \sum_{k=0}^{\infty} q^k = p \frac{1}{1-q} = 1.$$

$$E(X) = \sum_{k=0}^{\infty} k q^k p$$

$$E(X) = 0 \cdot p + (1 + E(X)) q$$

Strong proof: $c := E(X).$

$$E(X) = \frac{q}{p}$$

Negative Binomial (p.160, thm 4.3.9)

Sequence of $\text{Bern}(p)$.

$X = \#$ of failures before the r 'th success.

$$X \sim \text{NBin}(r, p)$$

Thm:

$$P(X=n) = \binom{n+r-1}{r-1} p^r q^n$$

Let X_i be # failures to first success.

$$X = X_1 + X_2 + \dots + X_r$$

$$E(X) = E(X_1 + X_2 + \dots + X_r) = E(X_1) + E(X_2) + \dots + E(X_r) = r \frac{q}{p}.$$

Ex Putnam problem (ex. 4.4.6)

$1, 2, 3, \dots, n$ permutation

4 2 5 1 3

$a_1, a_2, \dots, a_n$

when $n \geq 2$ what is the expected number of local maxima,
all permutations $n!$

I_j be indicator that there is local max at position j .

$$1 < j < n \quad \mathbb{E}(I_j) = \frac{1}{3}.$$

$$\mathbb{E}\left(\sum_{j=1}^n I_j\right) = 2 \cdot \frac{1}{3} + (n-2) \frac{1}{3} = \frac{n+1}{3}.$$

Lecture 6

• Functions of RV

- Variance
- Poisson
- Continuous
- Uniform

Functions of RV.

g is a function, taking values in $\mathbb{R}$.

$g(X)$

Suppose $E(g(X))$. Is this $g(E(X))$?

Thm. 4.5.1. Lotus

If X is discrete.

$g: \mathbb{R} \rightarrow \mathbb{R}$,

$$E(g(X)) = \sum_x g(x) P(g(X)=g(x)) = \sum_x g(x) P(X=x)$$

Proof

$$\begin{aligned} & \sum_s g(X(s)) P(\{s\}) \\ &= \sum_x \sum_{\substack{s: X(s)=x \\ \uparrow \\ \text{such that}}} g(X(s)) P(\{s\}) \\ &= \sum_x g(x) \underbrace{\sum_{s: X(s)=x} P(\{s\})}_{= P(X=x)} \end{aligned}$$

Motivation for Variance

$$\mathbb{E}(X - \mathbb{E}(X)) = 0 \quad \text{BAD}$$

$$\mathbb{E}(|X - \mathbb{E}(X)|) \neq \text{Better, but great.}$$

Def Variance 4.6.1

$$\text{VAR}(X) = \mathbb{E}((X - \mathbb{E}(X))^2)$$

Standard deviation

$$\text{SD}(X) = \sqrt{\text{VAR}(X)}$$

Suppose $\text{VAR}(X) = 0$. Then X is a constant.

VARIANCE IDENTITY (see 4.6)

$$\text{For any RV } X, \quad \text{VAR}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$$

PS: $g(x) = x^2$, a RV

$$\mu := \mathbb{E}(X)$$

$$\mathbb{E}(X - \mu)^2 = \mathbb{E}(X^2 - 2\mu X + \mu^2) = \mathbb{E}(X^2) - 2\mu \underbrace{\mathbb{E}(X)}_{\mu} + \mu^2$$

Facts about VARIANCE:

$$\text{VAR}(X + c) = \text{VAR}(X) \quad \text{for any constant } c$$

$$\text{VAR}(cX) = c^2 \text{VAR}(X)$$

If X and Y are independent, then

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

PS: Suppose $X = Y$.

$$\text{VAR}(X + Y) = \text{VAR}(2X) = 4 \text{VAR}(X) > 2 \text{VAR}(X)$$

Ex Binomial Variance (4.6.5)

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

$$X = I_1 + I_2 + \dots + I_n \quad I_j \text{ is an indicator of success in exp. } j.$$

$$\text{VAR}(I_j) = \mathbb{E}(I_j^2) - \mathbb{E}(I_j)^2 =$$

$$I_j^2 = I_j$$

$$= p - p^2 = p(1-p)$$

I_j are indep. Thus,

$$\text{VAR}(X) = \text{VAR}(I_1) + \text{VAR}(I_2) + \dots + \text{VAR}(I_n) = np(1-p)$$

Def 4.7.1 Poisson distribution

X has a poisson dist. with parameter λ , $X \sim \text{Pois}(\lambda)$, $\lambda > 0$,

if the PMF is
$$P(X=k) = \frac{e^{-\lambda} \lambda^k}{k!}, \quad k=0, 1, 2, \dots$$

$$\sum_k P(X=k) = \sum_{k=0}^{\infty} \frac{e^{-\lambda} \lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} e^{\lambda} = 1$$

Expectation 4.7.2

$$E(X) = e^{-\lambda} \sum_{k=0}^{\infty} \frac{k \lambda^k}{k!} = e^{-\lambda} \sum_{k=1}^{\infty} \frac{k \lambda^k}{k!} = \lambda e^{-\lambda} \underbrace{\sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!}}_{=e^{\lambda}} = \lambda.$$

$$\text{VAR}(X) = \lambda$$

Poisson paradigm 4.7

Intuition

Applications when we want to count something,
e.g. successes, when both

- Large number of trials
- success in each trial is small

Let $A_1, \dots, A_n$ be events with $p_j = P(A_j)$, p_j be small, A_j be weakly dependent

$$X = \sum_{j=1}^n \mathbb{I}(A_j),$$

then X is approx $\text{Pois}(\lambda)$, $\lambda = \sum_{j=1}^n p_j$

Thm 4.8.3.

$X \sim \text{Bin}(n, p)$, $n \rightarrow \infty$, $p \rightarrow 0$, $\lambda = np$ remain fixed.

Then the PMF of X converges to $\text{Pois}(\lambda)$.

Ex Birth triple match.

n people.

Find the approximate probability that 3 people (or more) have the same birthday

$\binom{n}{3}$ triplets.

I_{ijk} $i < j < k$ is an indicator

that i, j and k have the same birthday

I_{123} and I_{124} are dep. } "most triplets" are independent
 I_{123} and I_{456} are ind.

$$\begin{aligned} \mathbb{E}(\overbrace{\# \text{ number of triplets}}^{=X}) &= \binom{n}{3} \frac{1}{365^2}. \quad X \text{ is approx } \text{Pois}(\lambda), \quad \lambda = \binom{n}{3} \frac{1}{365^2}. \quad P(X \geq 1) = 1 - P(X=0) \\ &= 1 - e^{-\lambda} \frac{\lambda^0}{0!} = \underline{\underline{1 - e^{-\lambda}}}. \end{aligned}$$