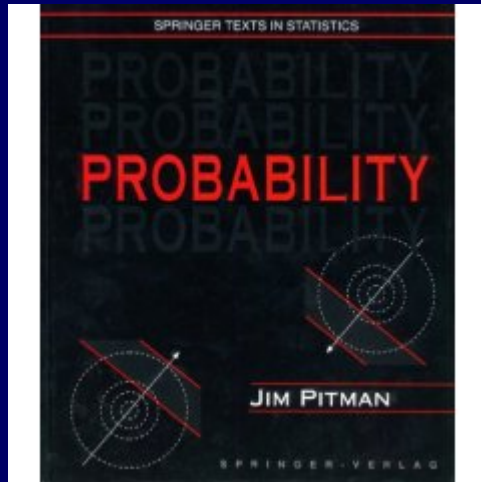


Introduction to probability

Stat 134

FALL 2005

Berkeley



Lectures prepared by:
Elchanan Mossel
Yelena Shvets

Follows Jim Pitman's
book:

Probability
Section 3.2

Multinomial Distribution

- Suppose each trial can result in m possible categories c_1, c_2, \dots, c_m with probabilities p_1, p_2, \dots, p_m , where $p_1 + p_2 + \dots + p_m = 1$.
- Suppose we make a sequence of n independent trials and let N_i denote the number of results in the i^{th} category c_i .

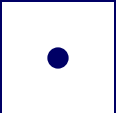
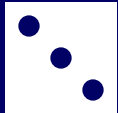
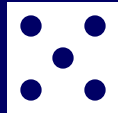
Multinomial Distribution

Then for every m -tuple of non-negative integers (n_1, n_2, \dots, n_m) with $n_1 + n_2 + \dots + n_m = n$

$$P(N_1 = n_1, N_2 = n_2, \dots, N_m = n_m) = \frac{n!}{n_1! n_2! \dots n_m!} p_1^{n_1} p_2^{n_2} \dots p_m^{n_m}$$

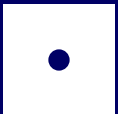
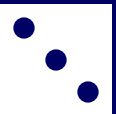
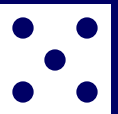
1, 3, 5 and 'even'

Suppose we roll a fair die 10 times and record the number of

 ,  ,  and even's.

Question:

What's the probability of seeing

1  , 2  , 3  and 4 even numbers?

1, 3, 5 and 'even'

Using the multinomial distribution:

$$P(N_1=1, N_3=2, N_5=3, N_{\text{even}}=4) = \frac{10!}{1!2!3!4!} \left(\frac{1}{6}\right)^1 \left(\frac{1}{6}\right)^2 \left(\frac{1}{6}\right)^3 \left(\frac{3}{6}\right)^4$$
$$= 0.016878858$$

Mean of a Distribution

- The **mean** μ of a probability distribution $P(x)$ over a finite set of numbers x
- Is the average of these numbers weighted by their probabilities:

$$\mu = \sum_x x P(X=x)$$

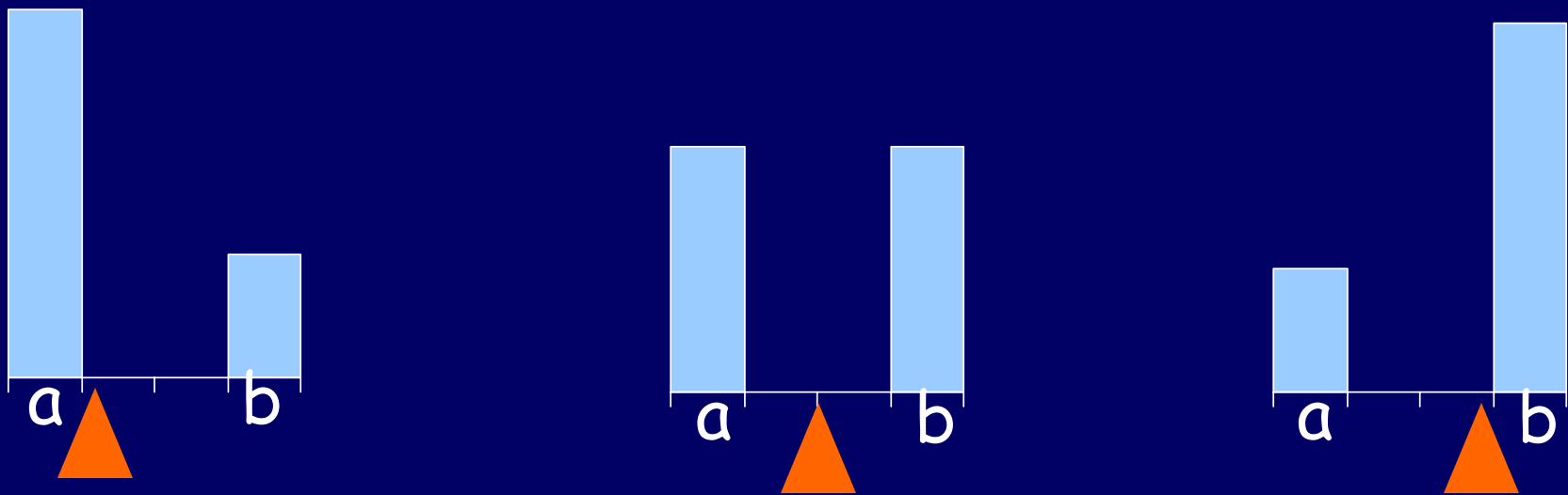
Expectation

The **expectation** (also **expected value** or **mean**) of a random variable X is the mean of the distribution of X .

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

Two Value Distributions

- If X is a Bernoulli(p) variable over $\{a, b\}$, then $E(X) = pa + (1-p)b$.
- If we think of p and q as two masses sitting over a and b then $E(X)$ would correspond to the point of balance:



A Fair Die

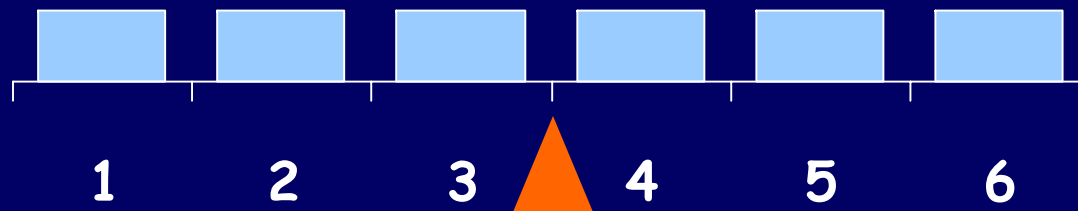
Let X be the number rolled with a fair die.

Question: What is the expected value of X ?

We can compute $E(X)$ by definition:

$$\begin{aligned} E(X) &= 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} \\ &= 3.5 \end{aligned}$$

Alternatively, we could find the point of balance on the histogram:



Binomial(10, $\frac{1}{2}$)

Question: Let Z be a variable with a binomial(10, $\frac{1}{2}$) distribution.

What is $E[Z]$?

By definition:

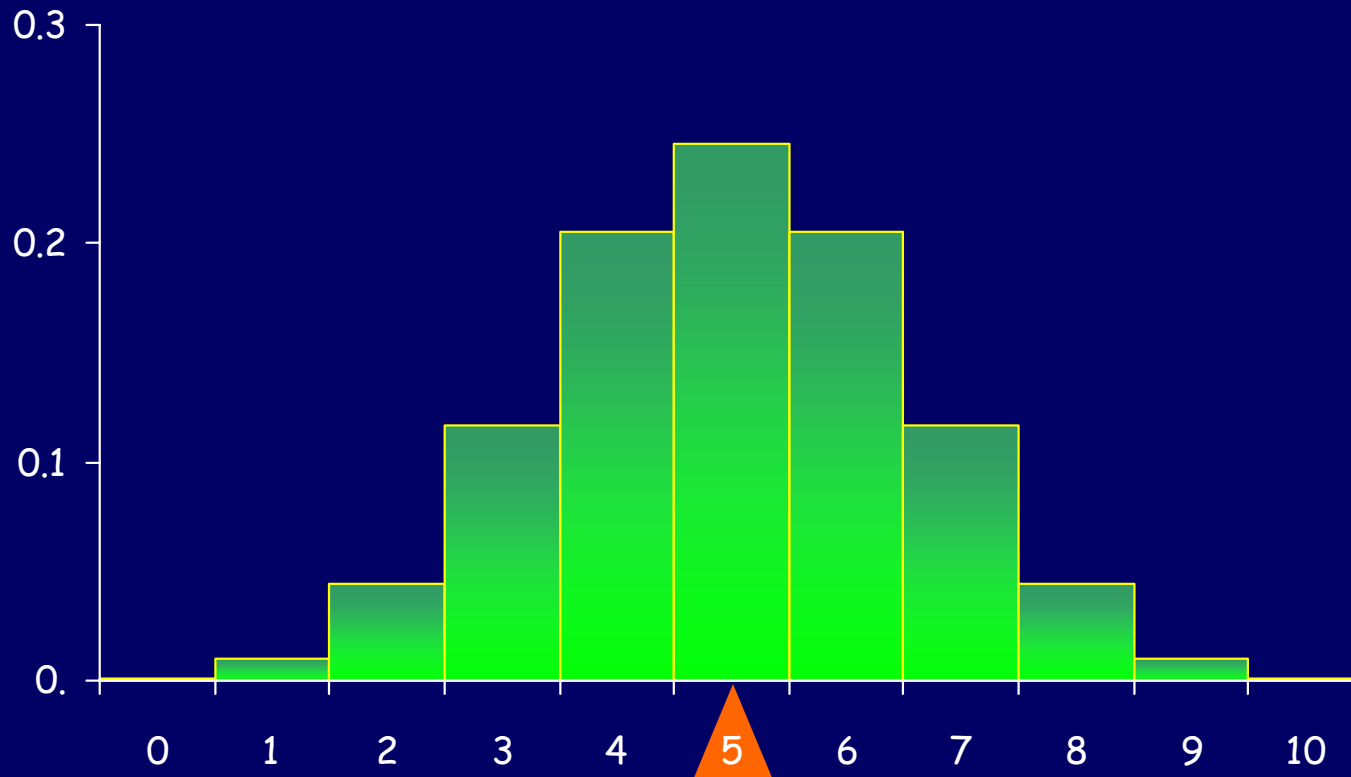
$$E(Z) = \sum_{i=0}^{10} iP(Z=i)$$

Binomial(10, $\frac{1}{2}$)

$$\begin{aligned} E(Z) &= 0 + 1 \cdot 10 \binom{10}{1} \left(\frac{1}{2}\right)^{10} + 2 \cdot 45 \binom{10}{2} \left(\frac{1}{2}\right)^{10} + 3 \cdot 120 \binom{10}{3} \left(\frac{1}{2}\right)^{10} + 4 \cdot 210 \binom{10}{4} \left(\frac{1}{2}\right)^{10} + 5 \cdot 252 \binom{10}{5} \left(\frac{1}{2}\right)^{10} \\ &\quad + 6 \cdot 210 \binom{10}{6} \left(\frac{1}{2}\right)^{10} + 7 \cdot 120 \binom{10}{7} \left(\frac{1}{2}\right)^{10} + 8 \cdot 45 \binom{10}{8} \left(\frac{1}{2}\right)^{10} + 9 \cdot 10 \binom{10}{9} \left(\frac{1}{2}\right)^{10} + 10 \binom{10}{10} \left(\frac{1}{2}\right)^{10} \\ &= \left(\frac{1}{2}\right)^{10} (10 + 90 + 360 + 840 + 1260 + 1260 + 840 + 360 + 90 + 10) \\ &= \frac{5120}{1024} = 5 \end{aligned}$$

Binomial(10, $\frac{1}{2}$)

We could also look at the histogram:



Addition Rule

• For any two random variables X and Y defined over the same sample space

$$E(X+Y) = E(X) + E(Y).$$

• Consequently, for a sequence of random variables X_1, X_2, \dots, X_n ,

$$E(X_1+X_2+\dots+X_n) = E(X_1) + E(X_2) + \dots + E(X_n).$$

Multiplication Rule and Inequalities

- Multiplication rule: $E[aX] = a E[X]$.
- $E[aX] = \sum_x a x P[X = x] = a \sum_x P[X = x] = a E[x]$
- If $X \geq Y$ then $E[X] \geq E[Y]$.
- This follows since $X - Y$ is non negative and
- $E[X] - E[Y] = E[X - Y] \geq 0$.

Sum of Two Dice

• Let T be the sum of two dice. What's $E(T)$?

• The "easy" way:

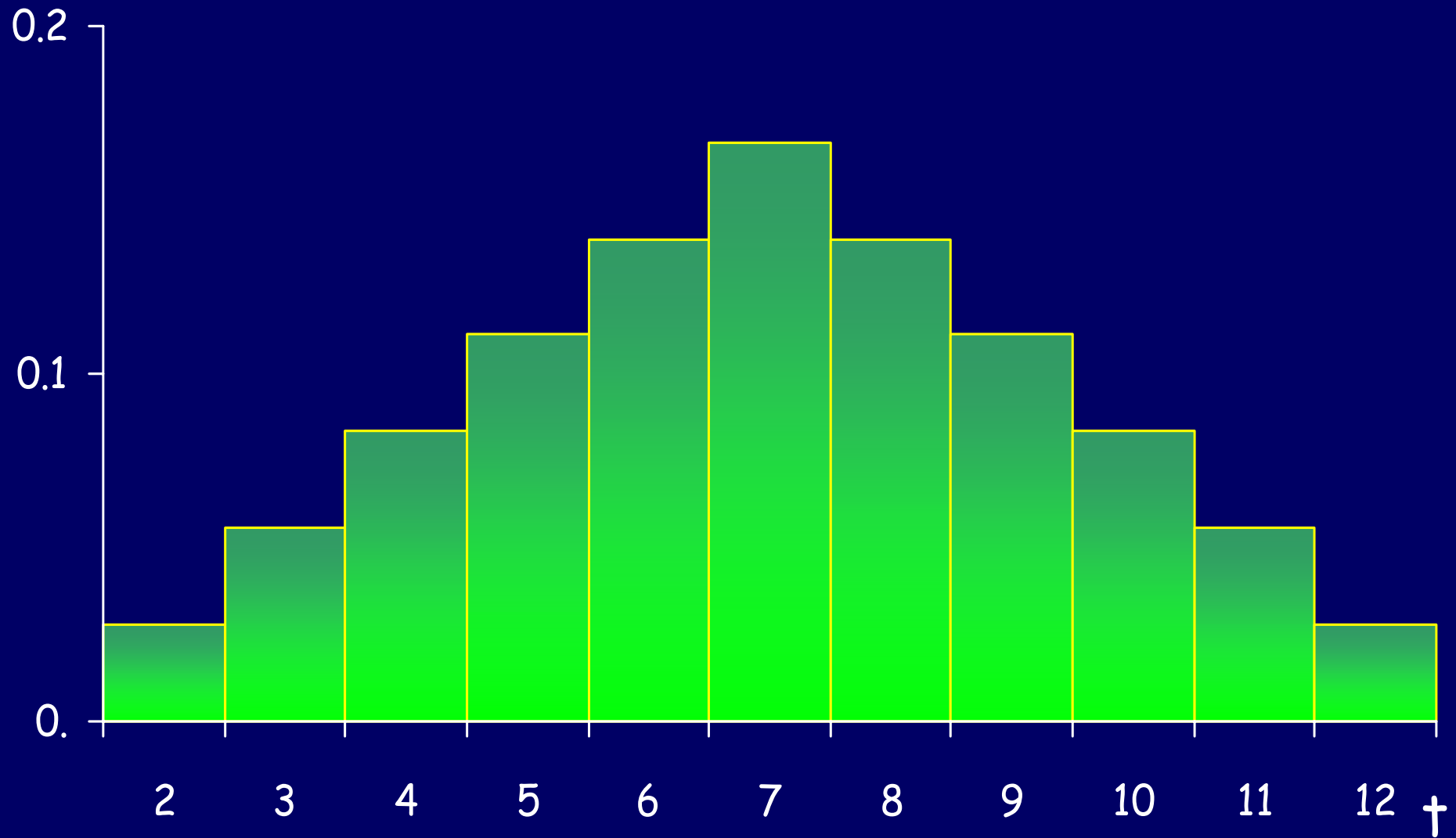
$$E(T) = \sum_{+} tP(T=t).$$

This sum will have 11 terms.

• We could also find the center of mass of the histogram (easy to do by symmetry).

Probability distribution for T.

$P(T=t)$



Sum of Two Dice

•Or we can using the addition rule:

$T = X_1 + X_2$, where $X_1 = 1^{\text{st}}$ role, $X_2 = 2^{\text{nd}}$:

$$E(T) = E(X_1) + E(X_2) = 3.5 + 3.5 = 7.$$

Indicators

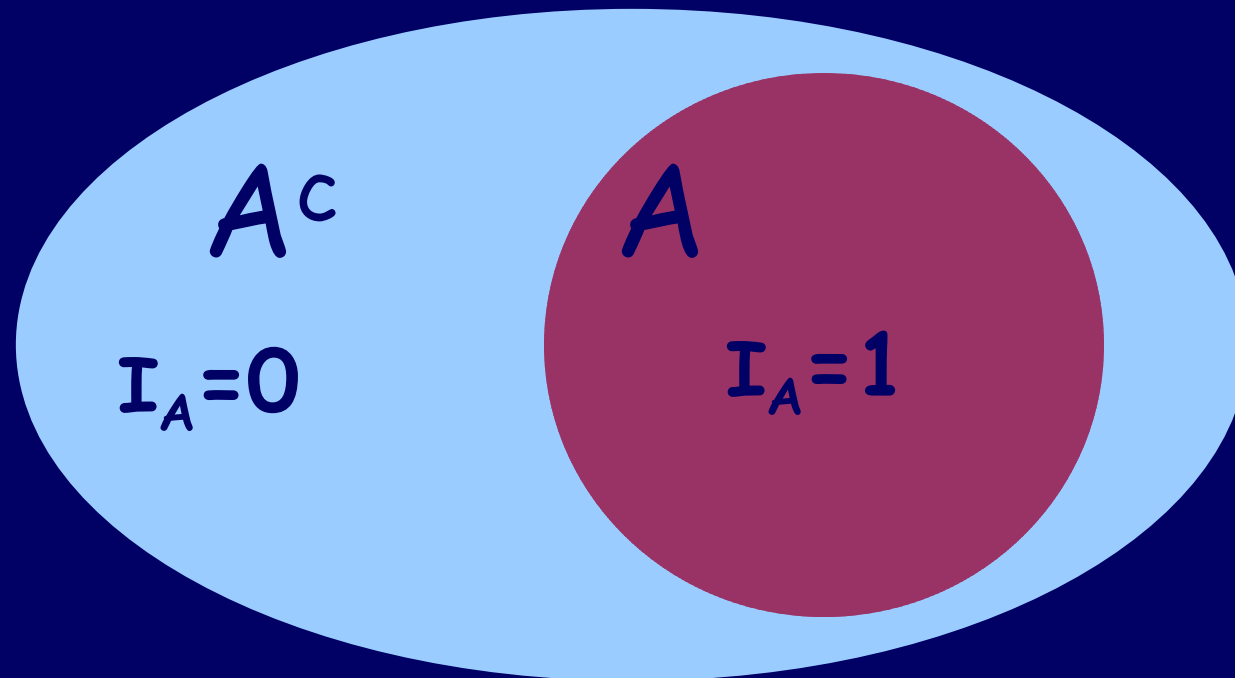
Indicators associate 0/1 valued random variables to events.

Definition: The **indicator** of the event A ,

I_A is the random variable that takes the value 1 for outcomes in A and the value 0 for outcomes in A^c .

Indicators

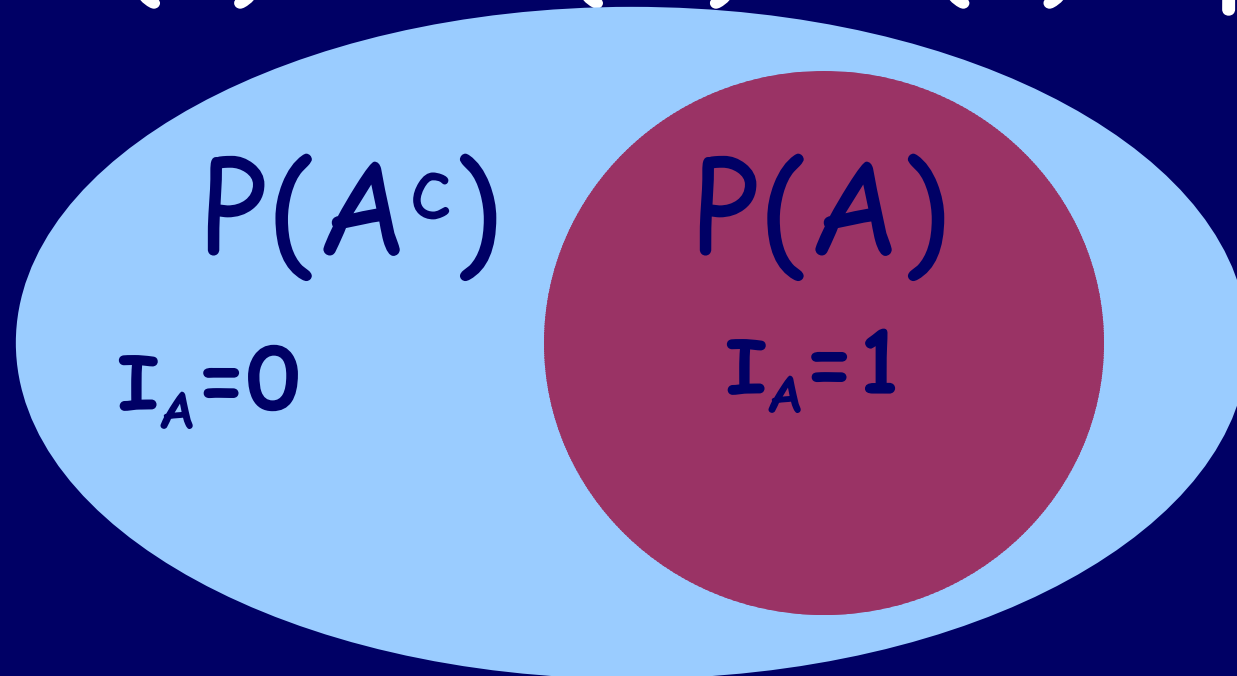
Suppose I_A is an indicator of an event A with probability p .



Expectation of Indicators

Then:

$$E(I_A) = 1 * P(A) + 0 * P(A^c) = P(A) = p.$$



Expected Number of Events that Occur

- Suppose there are n events A_1, A_2, \dots, A_n .
- Let $X = I_1 + I_2 + \dots + I_n$ where I_i is the indicator of A_i .
- Then X counts the number of events that occur.
- By the addition rule:

$$E(X) = P(A_1) + P(A_2) + \dots + P(A_n).$$

Repeated Trials

- Let X^i = **indicator** of success on the i^{th} coin toss ($X^i = 1$ if the i^{th} coin toss = H head, and $X^i = 0$ otherwise).
- The sequence X^1, X^2, \dots, X^n is a sequence of n independent variables with Bernoulli(p) distribution over $\{0,1\}$.
- The number of heads in n coin tosses given by $S_n = X^1 + X^2 + \dots + X^n$.
- $E(S_n) = nE(X^i) = np$
- Thus the mean of Bin(n,p) RV = np .

Expected Number of Aces

- Let Y be the number of aces in a poker hand.
- Then:

$$Y = I_{1\text{st ace}} + I_{3\text{rd ace}} + I_{4\text{th ace}} + I_{5\text{th ace}} + I_{2\text{nd ace}}.$$

- And: $E(Y) = 5 * P(\text{ace}) = 5 * 4/52 = 0.385$.
- Alternatively, since Y has the hypergeometric distribution we can calculate:

$$E(Y) = \sum_{x=0}^4 y \frac{\binom{4}{y} \binom{48}{5-y}}{\binom{52}{5}}$$

Non-negative Integer Valued RV

- Suppose X is an integer valued, non-negative random variable.
- Let $A_i = \{X \geq i\}$ for $i=1,2,\dots$;
- Let I_i the indicator of the set A_i .
- Then

$$X = \sum_i I_i.$$

Non-negative Integer Valued RV

• The equality

$$X(\text{outcome}) = \sum_i I_i(\text{outcome})$$

follows since if $X(\text{outcome}) = i$, then

outcome $\in A_1 \cap A_2 \cap \dots \cap A_i$ but not to $A_j, j > i$.

• So $(I_1 + I_2 + \dots + I_i + I_{i+1} + \dots)(\text{outcome}) =$

$$1 + 1 + \dots + 1 + 0 + 0 + \dots = i.$$

Tail Formula for Expectation

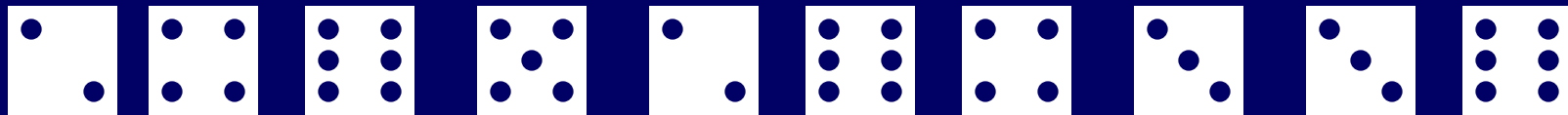
Let X be a non-negative integer valued RV,

Then:

$$\begin{aligned} E(X) &= E(\sum_i I_i) = \sum_i E(I_i) \\ &= \sum_i P(X \geq i), \quad i=1,2,3\dots \end{aligned}$$

Minimum of 10 Dice

Suppose we roll a die 10 times and let X be the minimum of the numbers rolled.



Here $X = 1$.

Question: what's the expected value of X ?

Minimum of 10 Dice

• Let's use the tail formula to compute $E(X)$:

$$E(X) = \sum_i P(X \geq i).$$

$$P(X \geq 1) = 1;$$

$$P(X \geq 2) = (5/6)^{10};$$

$$P(X \geq 3) = (4/6)^{10};$$

$$P(X \geq 4) = (3/6)^{10};$$

$$P(X \geq 5) = (2/6)^{10};$$

$$P(X \geq 6) = (1/6)^{10}$$

$$\begin{aligned} E(X) &= (6^{10} + 5^{10} + 4^{10} + 3^{10} + 2^{10} + 1^{10}) / 6^{10} \\ &= 1.17984 \end{aligned}$$

Indicators

• If the events A_1, A_2, \dots, A_j are mutually exclusive then

$$I_1 + I_2 + \dots + I_j = I_{A_1 \cup A_2 \cup \dots \cup A_j}$$

• And

$$P(\cup_{i=1}^j A_i) = \sum_i P(A_i).$$

Tail Formula for Expectation

Let X be a non-negative integer valued RV,

Then:

$$\begin{aligned} E(X) &= E(\sum_i I_i) = \sum_i E(I_i) \\ &= \sum_i P(X \geq i), \quad i=1,2,3\dots \end{aligned}$$

Boole's Inequality

For a non-negative integer valued X we can obtain **Boole's** inequality:

$$P(X \geq 1) \leq \sum_i P(X \geq i) = E(X)$$

Markov's Inequality

Markov inequality:

If $X \geq 0$, then for every $a > 0$

$$P(X \geq a) \leq E(X)/a.$$

- This is proven as follows.
- Note that if $X \leq Y$ then $E(X) \leq E(Y)$.
- Take $Y =$ indicator of the event $\{X \geq a\}$.
- Then $E(Y) = a P(X \geq a)$ and $X \leq aY$ so:
- $E(X) \leq E(aY) = a E(Y) = a P(X \geq a)$.

Expectation of a Function of a Random Variable

- For any function g defined on the range space of a random variable X with a finite number of values

$$E[g(X)] = \sum_x g(x) P(X=x).$$

Proof:

- Note that:

$$P(g(X)=y) = \sum_{\{x:g(x)=y\}} P(X=x).$$

- Therefore:

$$\begin{aligned} E[g(X)] &= \sum_y y P(g(X)=y) = \sum_y \sum_{\{x:g(x)=y\}} g(x) P(X=x) \\ &= \sum_x g(x) P(X=x). \end{aligned}$$

Expectation of a Function of a Random Variable

- Constants:

$$g(X)=c \Rightarrow E[g(x)]=c.$$

- Linear functions:

$$g(X)=aX + b \Rightarrow E[g(x)]=aE(X)+b.$$

(These are the only cases when $E(g(X)) = g(E(X))$.)

Expectation of a Function of a Random Variable

- Monomials:

$$g(X)=X^k \Rightarrow E[g(x)]=\sum_x x^k P(X=x).$$

$\sum_x x^k P(X=x)$ is called the k^{th} moment of X .

Expectation of a Function of a Random Variable

Question: For X representing the number on a die, what is the **second moment** of X ?

$$\begin{aligned}\sum_x x^2 P(X=x) &= \sum_x x^2 / 6 = 1/6 * (1 + 4 + 9 + 16 + 25 + 36) \\ &= 91/6 = 15.16667\end{aligned}$$

Expectation of a Function of Several Random Variables

• If X and Y are two random variables we obtain:

$$E(g(X,Y)) = \sum_{\{\text{all } (x,y)\}} g(x,y)P(X=x, Y=y).$$

This allows to prove that $E[X+Y] = E[X] + E[Y]$:

$$E(X) = \sum_{\{\text{all } (x,y)\}} x P(X=x, Y=y);$$

$$E(Y) = \sum_{\{\text{all } (x,y)\}} y P(X=x, Y=y);$$

$$E(X+Y) = \sum_{\{\text{all } (x,y)\}} (x+y) P(X=x, Y=y);$$

$$E(X+Y) = E(X) + E(Y)$$

Expectation of a Function of Several Random Variables

$$E(g(X,Y)) = \sum_{\{\text{all } (x,y)\}} g(x,y)P(X=x, Y=y).$$

Product:

$$E(XY) = \sum_{\{\text{all } (x,y)\}} xy P(X=x, Y=y);$$

$$E(XY) = \sum_x \sum_y xy P(X=x, Y=y);$$

Typically $E(XY) \neq E(X) E(Y)$.

Product Rule for Independent Random Variables

- However, if X and Y are **independent**,

$$P(X=x, Y=y) = P(X=x)P(Y=y)$$

then product formula simplifies:

$$\begin{aligned} E(XY) &= \sum_x \sum_y xy P(X=x) P(Y=y) \\ &= (\sum_x x P(X=x)) (\sum_y y P(Y=y)) = \\ &E(X) E(Y) \end{aligned}$$

- If X and Y are independent then:

$$E(XY) = E(X) E(Y);$$

Roulette



		0	00
1-18	I	1	2
Even	J	4	5
Red	K	7	8
Black	L	10	11
Odd	M	13	14
19-36	N	16	17
		19	20
		22	23
		25	26
		28	29
		31	32
		34	35
		36	
		G	2-1

Bet

- Straight Up (one number)
- Split
- Line/Street (three numbers)
- Corner (four numbers)
- Five line (5 numbers-0,00,1,2 or 3)
- Six line (six numbers)
- Column (twelve numbers)
- Dozens (twelve numbers)
- 1 - 18
- Even
- Red
- Black
- Odd
- 19-36

Pay-off

- 35:1
- 17:1
- 11:1
- 8:1
- 6:1
- 5:1
- 2:1
- 2:1
- 1:1
- 1:1
- 1:1
- 1:1
- 1:1
- 1:1

Betting on Red

Suppose we want to bet \$1 on Red. Our chance of winning is $18/38$.

Question:

What should be the pay-off to make it a fair bet?

Betting on Red

This question really only makes sense if we repeatedly bet \$1 on Red.

Suppose that we could win \$ x if Reds come up and lose \$1, otherwise. If X denotes our returns then $P(X=x) = 18/38$; $P(X=-1)=20/38$.

In a **fair game**, we expect to **break even** on average.

Betting on Red

Our expected return is:

$$x \cdot 18/38 - 1 \cdot 20/38.$$

Setting this to zero gives us

$$x = 20/18 = 1.1111111... .$$

This is greater than the pay-off of 1:1 that is offered by the casino.

Expectation interpretation as a Long-Run Average

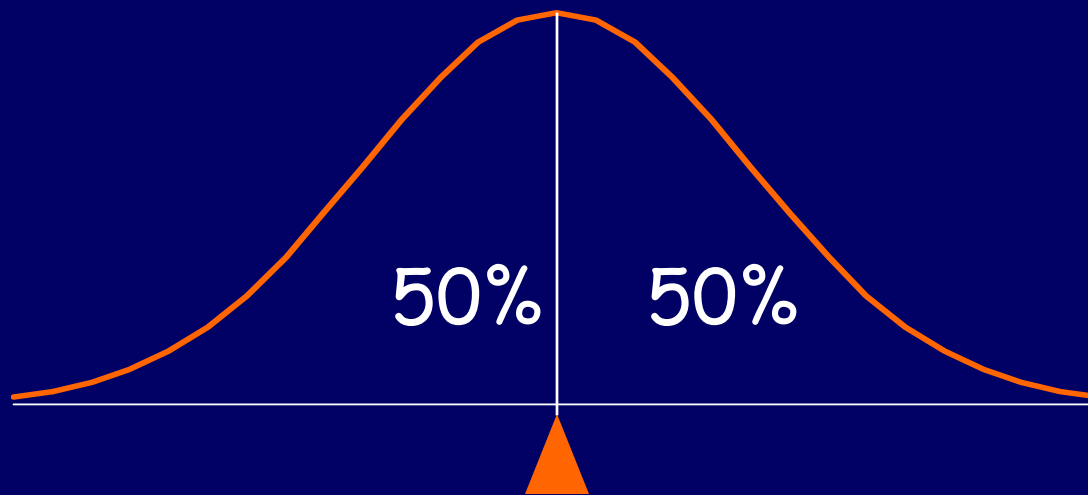
- If we repeatedly sample from the distribution of X then $P(X=x)$ will be close to the observed frequency of x in the sample.
- $E(X)$ will be approximately the long-run average of the sample.

Mean, Mode and Median

- The **Mode** of X is the most likely possible value of X .
- The mode need not be unique.
- The **Median** of X is a number m such that both $P(X \leq m) \geq \frac{1}{2}$ and $P(X \geq m) \geq \frac{1}{2}$.
- The median may also not be unique.
- **Mean** and **Median** are not necessarily possible values (mode is).

Mean, Mode and Median

For a symmetrical distribution, which has a unique *Mode*, all three: *Mean*, *Mode* and *Median* are the same.



mean = mode = median

Mean, Mode and Median

For a distribution with a long right tail Mean is greater than the Median.

