CS70: Jean Walrand: Lecture 18.

Random Variables & Midterm 2 Probability Review

- Random Variables
- M2 Probability Review
- M2 Discrete Math Review: See Video (link given on Piazza)

Random Variables

1. Random Variables.
2. Distributions.
3. Combining random variables.
4. Expectation

Questions about outcomes ...

Experiment: roll two dice.
Sample Space: $\{(1,1),(1,2),\dots,(6,6)\} = \{1,\dots,6\}^2$
How many pips?

Experiment: flip 100 coins.
Sample Space: $\{HHH\cdots H, THH\cdots H, \dots, TTT\cdots T\}$
How many heads in 100 coin tosses?

Experiment: choose a random student in cs70.
Sample Space: $\{Adam, Jin, Bing, \dots, Angeline\}$
What midterm score?

Experiment: hand back assignments to 3 students at random.
Sample Space: $\{123, 132, 213, 231, 312, 321\}$
How many students get back their own assignment?

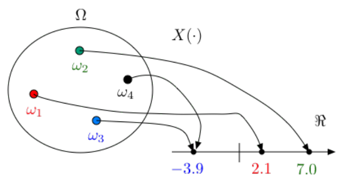
In each scenario, each outcome gives a number.

The number is a (known) function of the outcome.

Random Variables.

A **random variable,** X**,** for an experiment with sample space Ω is a function $X : \Omega \to \Re$.

Thus, $X(\cdot)$ assigns a real number $X(\omega)$ to each $\omega \in \Omega$.



$X(\omega_1) = 2.1; X(\omega_2) = 7.0; X(\omega_3) = X(\omega_4) = -3.9$

The function $X(\cdot)$ is defined on the outcomes Ω.

The function $X(\cdot)$ is not random, not a variable!

What varies at random (from experiment to experiment)? The outcome!

Example 1 of Random Variable

Experiment: roll two dice.
Sample Space: $\{(1,1),(1,2),\dots,(6,6)\} = \{1,\dots,6\}^2$
Random Variable X: number of pips.
$X(1,1) = 2$
$X(1,2) = 3$,
$\vdots$
$X(6,6) = 12$,
$X(a,b) = a + b, (a,b) \in \Omega$.

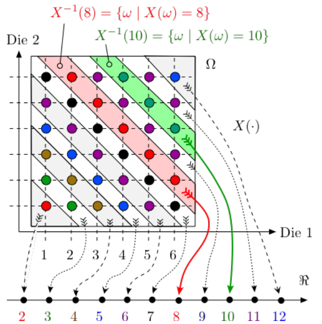
Example 2 of Random Variable

Experiment: flip three coins
Sample Space: $\{HHH, THH, HTH, TTH, HHT, THT, HTT, TTT\}$
Winnings: if win 1 on heads, lose 1 on tails: X

$X(HHH) = 3$ $X(THH) = 1$ $X(HTH) = 1$ $X(TTH) = -1$
$X(HHT) = 1$ $X(THT) = -1$ $X(HTT) = -1$ $X(TTT) = -3$

Number of pips in two dice.

"What is the likelihood of getting n pips?"

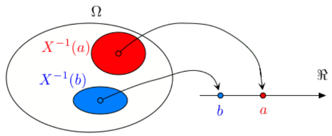


$Pr[X=10] = 3/36 = Pr[X^{-1}(10)]$; $Pr[X=8] = 5/36 = Pr[X^{-1}(8)]$.

Distribution

The probability of X taking on a value a.

Definition: The **distribution** of a random variable X, is $\{(a, Pr[X=a]) : a \in \mathscr{A}\}$, where $\mathscr{A}$ is the range of X.



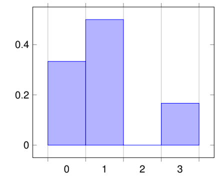
$Pr[X=a] := Pr[X^{-1}(a)]$ where $X^{-1}(a) := \{\omega \mid X(\omega) = a\}$.

Handing back assignments

Experiment: hand back assignments to 3 students at random.
Sample Space: $\Omega = \{123, 132, 213, 231, 312, 321\}$
How many students get back their own assignment?
Random Variable: values of $X(\omega) : \{3,1,1,0,0,1\}$

Distribution:

$$X = \begin{cases} 0, & \text{w.p. } 1/3 \\ 1, & \text{w.p. } 1/2 \\ 3, & \text{w.p. } 1/6 \end{cases}$$

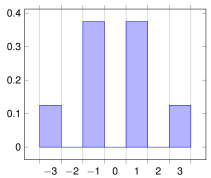


Flip three coins

Experiment: flip three coins
Sample Space: $\{HHH, THH, HTH, TTH, HHT, THT, HTT, TTT\}$
Winnings: if win 1 on heads, lose 1 on tails. X
Random Variable: $\{3,1,1,-1,1,-1,-1,-3\}$

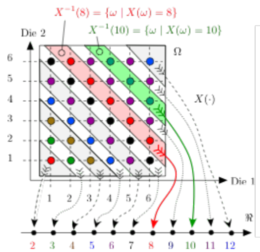
Distribution:

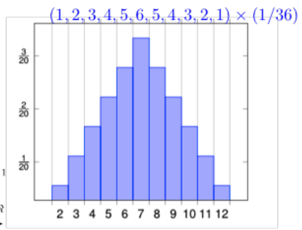
$$X = \begin{cases} -3, & \text{w. p. } 1/8 \\ -1, & \text{w. p. } 3/8 \\ 1, & \text{w. p. } 3/8 \\ 3 & \text{w. p. } 1/8 \end{cases}$$



Number of pips.

Experiment: roll two dice.





The binomial distribution.

Flip n coins with heads probability p.

Random variable: number of heads.

Binomial Distribution: $Pr[X=i]$, for each i.

How many sample points in event "$X=i$"?
i heads out of n coin flips $\Longrightarrow \binom{n}{i}$

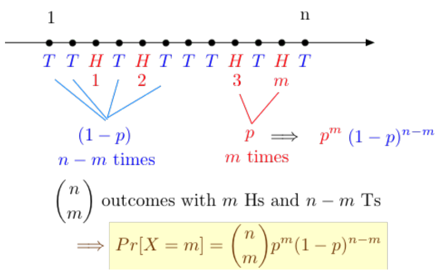
What is the probability of ω if ω has i heads?
Probability of heads in any position is p.
Probability of tails in any position is $(1-p)$.
So, we get

$$Pr[\omega] = p^i (1-p)^{n-i}.$$

Probability of "$X=i$" is sum of $Pr[\omega]$, $\omega \in$ "$X=i$".

$$Pr[X=i] = \binom{n}{i} p^i (1-p)^{n-i}, i = 0,1,\ldots,n : B(n,p) \text{ distribution}$$

The binomial distribution.



Error channel.

A packet is corrupted with probability p.

Send $n+2k$ packets.

Probability of at most k corruptions.

$$\sum_{i\le k}\binom{n+2k}{i}p^i(1-p)^{n+2k-i}.$$

Confidence in polling, experiments, etc.

Combining Random Variables.

Let X and Y be two RV on the same probability space.

That is, $X:\Omega\to\Re$ assigns the value $X(\omega)$ to ω. Also, $Y:\Omega\to\Re$ assigns the value $Y(\omega)$ to ω.

Then $X+Y$ is a random variable: It assigns the value

$$X(\omega)+Y(\omega)$$

to ω.

Experiment: Roll two dice. X = outcome of first die, Y = outcome of second die. Thus,

$$X(a,b)=a \text{ and } Y(a,b)=b \text{ for } (a,b)\in\Omega=\{1,\dots,6\}^2.$$

Then $Z=X+Y$ = sum of two dice is defined by

$$Z(a,b)=X(a,b)+Y(a,b)=a+b.$$

Combining Random Variables

Other random variables:

- $X^k:\Omega\to\Re$ is defined by $X^k(\omega)=[X(\omega)]^k$. In the dice example, $X^3(a,b)=a^3$.
- $(X-2)^2+4XY$ assigns the value $(X(\omega)-2)^2+4X(\omega)Y(\omega)$ to ω.
- $g(X,Y,Z)$ assigned the value $g(X(\omega),Y(\omega),Z(\omega))$ to ω.

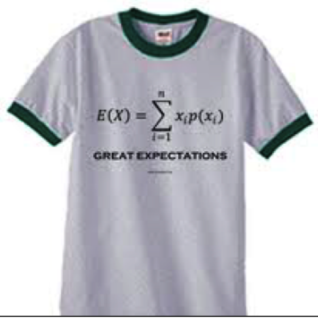
Expectation.

How did people do on the midterm?

Distribution.

Summary of distribution?

Average!



Expectation - Intuition

Flip a loaded coin with $Pr[H]=p$ a large number N of times.

We expect heads to come up a fraction p of the times and tails a fraction $1-p$.

Say that you get 5 for every H and 3 for every T.

If there are $N(H)$ outcomes equal to H and $N(T)$ outcomes equal to T, you collect

$$5\times N(H)+3\times N(T).$$

pause You average gain per experiment is then

$$\frac{5N(H)+3N(T)}{N}.$$

Since $\frac{N(H)}{N}\approx p=Pr[X=5]$ and $\frac{N(T)}{N}\approx 1-p=Pr[X=3]$, we find that the average gain per outcome is approximately equal to

$$5Pr[X=5]+3Pr[X=3].$$

We use this frequentist interpretation as a definition.

Expectation - Definition

Definition: The **expected value** of a random variable X is

$$E[X] = \sum_a a \times Pr[X = a].$$

The expected value is also called the mean.

According to our intuition, we expect that if we repeat an experiment a large number N of times and if $X_1, \ldots, X_N$ are the successive values of the random variable, then

$$\frac{X_1 + \cdots + X_N}{N} \approx E[X].$$

That is indeed the case, in the same way that the fraction of times that $X = x$ approaches $Pr[X = x]$.

This (nontrivial) result is called the Law of Large Numbers.

The subjectivist interpretation of $E[X]$ is less obvious.

Expectation: A Useful Fact

Theorem:

$$E[X] = \sum_\omega X(\omega) \times Pr[\omega].$$

Proof:

$$\begin{aligned} E[X] &= \sum_a a \times Pr[X = a] \\ &= \sum_a a \times \sum_{\omega : X(\omega) = a} Pr[\omega] \\ &= \sum_a \sum_{\omega : X(\omega) = a} X(\omega) Pr[\omega] \\ &= \sum_\omega X(\omega) Pr[\omega] \end{aligned}$$

An Example

Flip a fair coin three times.

$\Omega = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$.

$X =$ number of H's: $\{3, 2, 2, 2, 1, 1, 1, 0\}$.

Thus,

$$\sum_\omega X(\omega) Pr[\omega] = \{3 + 2 + 2 + 2 + 1 + 1 + 1 + 0\} \times \frac{1}{8}.$$

Also,

$$\sum_a a \times Pr[X = a] = 3 \times \frac{1}{8} + 2 \times \frac{3}{8} + 1 \times \frac{3}{8} + 0 \times \frac{1}{8}.$$

Expectation and Average.

There are n students in the class;

$X(m)$ = score of student m, for $m = 1, 2, \ldots, n$.

"Average score" of the n students: add scores and divide by n:

$$\text{Average} = \frac{X(1) + X(1) + \cdots + X(n)}{n}.$$

Experiment: choose a student uniformly at random.
Uniform sample space: $\Omega = \{1, 2, \cdots, n\}, Pr[\omega] = 1/n$, for all ω.
Random Variable: midterm score: $X(\omega)$.
Expectation:

$$E(X) = \sum_\omega X(\omega) Pr[\omega] = \sum_\omega X(\omega) \frac{1}{n}.$$

Hence,

$$\text{Average } = E(X).$$

This holds for a uniform probability space.

Handing back assignments

We give back assignments randomly to three students.
What is the expected number of students that get their own assignment back?

"The expected number of **fixed points** in a random permutation."
Expected value of a random variable:

$$E[X] = \sum_a a \times Pr[X = a].$$

For 3 students (permutations of 3 elements):

$$Pr[X = 3] = 1/6, Pr[X = 1] = 1/2, Pr[X = 0] = 1/3.$$

$$E[X] = 3 \times \frac{1}{6} + 1 \times \frac{1}{2} + 0 \times \frac{1}{3} = 1.$$

Win or Lose.

Expected winnings for heads/tails games, with 3 flips?

$$E[X] = 3 \times \frac{1}{8} + 1 \times \frac{3}{8} - 1 \times \frac{3}{8} - 3 \times \frac{1}{8} = 0.$$

Can you ever win 0?

Apparently: expected value is not a common value, by any means.

Expectation

Recall: $X : \Omega \to \Re; Pr[X = a] := Pr[X^{-1}(a)]$;

Definition: The **expectation** of a random variable X is

$$E[X] = \sum_a a \times Pr[X = a].$$

Indicator:

Let A be an event. The random variable X defined by

$$X(\omega) = \begin{cases} 1, & \text{if } \omega \in A \\ 0, & \text{if } \omega \notin A \end{cases}$$

is called the indicator of the event A.

Note that $Pr[X = 1] = Pr[A]$ and $Pr[X = 0] = 1 - Pr[A]$.

Hence,

$$E[X] = 1 \times Pr[X = 1] + 0 \times Pr[X = 0] = Pr[A].$$

The random variable X is sometimes written as

$$1\{\omega \in A\} \text{ or } 1_A(\omega).$$

Linearity of Expectation

Theorem:

$$E[X] = \sum_\omega X(\omega) \times Pr[\omega].$$

Theorem: Expectation is linear

$$E[a_1X_1 + \cdots + a_nX_n] = a_1E[X_1] + \cdots + a_nE[X_n].$$

Proof:

$$\begin{aligned} & E[a_1X_1 + \cdots + a_nX_n] \\ &= \sum_\omega (a_1X_1 + \cdots + a_nX_n)(\omega)Pr[\omega] \\ &= \sum_\omega (a_1X_1(\omega) + \cdots + a_nX_n(\omega))Pr[\omega] \\ &= a_1\sum_\omega X_1(\omega)Pr[\omega] + \cdots + a_n\sum_\omega X_n(\omega)Pr[\omega] \\ &= a_1E[X_1] + \cdots + a_nE[X_n]. \end{aligned}$$

□

Using Linearity - 1: Pips on dice

Roll a die n times.

X_m = number of pips on roll m.

$X = X_1 + \cdots + X_n$ = total number of pips in n rolls.

$$\begin{aligned} E[X] &= E[X_1 + \cdots + X_n] \\ &= E[X_1] + \cdots + E[X_n], \text{ by linearity} \\ &= nE[X_1], \text{ because the } X_m \text{ have the same distribution} \end{aligned}$$

Now,

$$E[X_1] = 1 \times \frac{1}{6} + \cdots + 6 \times \frac{1}{6} = \frac{6 \times 7}{2} \times \frac{1}{6} = \frac{7}{2}.$$

Hence,

$$E[X] = \frac{7n}{2}.$$

Using Linearity - 2: Fixed point.

Hand out assignments at random to n students.

X = number of students that get their own assignment back.

$X = X_1 + \cdots + X_n$ where

$X_m = 1\{\text{student } m \text{ gets his/her own assignment back}\}$.

One has

$$\begin{aligned} E[X] &= E[X_1 + \cdots + X_n] \\ &= E[X_1] + \cdots + E[X_n], \text{ by linearity} \\ &= nE[X_1], \text{ because all the } X_m \text{ have the same distribution} \\ &= nPr[X_1 = 1], \text{ because } X_1 \text{ is an indicator} \\ &= n(1/n), \text{ because student 1 is equally likely} \\ &\qquad \text{to get any one of the } n \text{ assignments} \\ &= 1. \end{aligned}$$

Note that linearity holds even though the X_m are not independent (whatever that means).

Using Linearity - 3: Binomial Distribution.

Flip n coins with heads probability p. X - number of heads

Binomial Distibution: $Pr[X = i]$, for each i.

$$Pr[X = i] = \binom{n}{i} p^i (1-p)^{n-i}.$$

$$E[X] = \sum_i i \times Pr[X = i] = \sum_i i \times \binom{n}{i} p^i (1-p)^{n-i}.$$

Uh oh. ... Or... a better approach: Let

$$X_i = \begin{cases} 1 & \text{if } i\text{th flip is heads} \\ 0 & \text{otherwise} \end{cases}$$

$E[X_i] = 1 \times Pr[\text{"heads"}] + 0 \times Pr[\text{"tails"}] = p.$

Moreover $X = X_1 + \cdots X_n$ and

$E[X] = E[X_1] + E[X_2] + \cdots E[X_n] = n \times E[X_i] = np.$

Summary

Random Variables

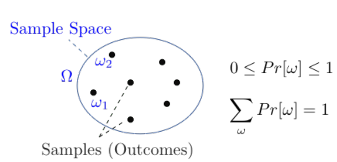
- A random variable X is a function $X : \Omega \to \Re$.
- $Pr[X = a] := Pr[X^{-1}(a)] = Pr[\{\omega \mid X(\omega) = a\}]$.
- $Pr[X \in A] := Pr[X^{-1}(A)]$.
- The distribution of X is the list of possible values and their probability: $\{(a, Pr[X = a]), a \in \mathscr{A}\}$.
- $g(X, Y, Z)$ assigns the value .... .
- $E[X] := \sum_a aPr[X = a]$.
- Expectation is Linear.
- $B(n, p)$.

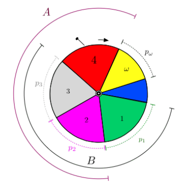
Probability: Midterm 2 Review.

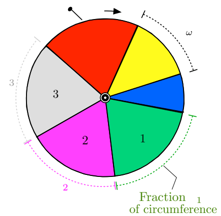
- Framework:
  - Probability Space
  - Conditional Probability & Bayes' Rule
  - Independence
  - Mutual Independence
- Collisions & Collecting
- Random Variables

See Note 25: 1, 2, 3, 4 (paragraphs 1, 2, 3; examples 1 through 8)

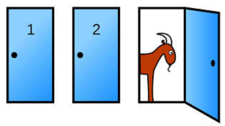
Probability Space





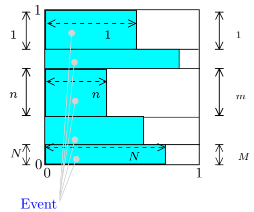


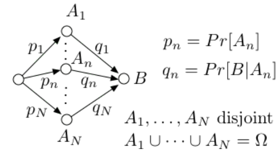
$Pr[A|B] = Pr[A \cap B]/Pr[B]$.
$Pr[A \cap B \cap C]$
$= Pr[A]Pr[B|A]Pr[C|A \cap B]$.



Bayes' Rule

- Priors: $Pr[A_n] = p_n, n = 1, \ldots, M$
- Conditional Probabilities: $Pr[B|A_n] = q_n, n = 1, \ldots, N$
- $\Rightarrow$ Posteriors: $Pr[A_n|B] = \frac{p_n q_n}{p_1 q_1 + \cdots + p_N q_N}$





Bayes' Rule: Examples

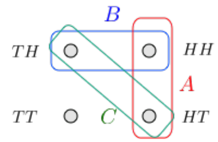
Let $p'_n = Pr[A_n|B]$ be the posterior probabilities.

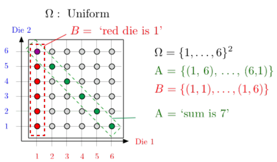
Thus, $p'_n = p_n q_n/(p_1 q_1 + \cdots + p_N q_n)$.

Questions: Is it true that

- if $q_n > q_k$, then $p'_n > p'_k$? Not necessarily.
- if $p_n > p_k$, then $p'_n > p'_k$? Not necessarily.
- if $p_n > p_k$ and $q_n > q_k$, then $p'_n > p'_k$? Yes.
- if $q_n = 1$, then $p'_n > 0$? Not necessarily.
- if $p_n = 1/N$ for all n, then MLE = MAP? Yes.

Independence





"First coin yields 1" and "Sum is 7" are independent

Pairwise, but not mutually

If $\{A_j, i \in J\}$ are mutually independent, then $[A_1 \cap \bar{A}_2] \Delta A_3$ and $A_4 \setminus A_5$ are independent.

Our intuitive meaning of "independent events" is mutual independence.

Independence

Recall

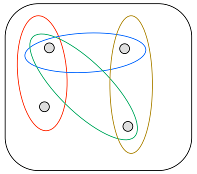
- A and B are independent if $Pr[A \cap B] = Pr[A]Pr[B]$.
- $\{A_j, j \in J\}$ are mutually independent if $Pr[\cap_{j \in K} A_j] = \Pi_{j \in K} Pr[A_j], \forall$ finite $K \subset J$.

Thus, A, B, C, D are mutually independent if there are

- independent 2 by 2: $Pr[A \cap B] = Pr[A]Pr[B], \ldots, Pr[C \cap D] = Pr[C]Pr[D]$
- by 3: $Pr[A \cap B \cap C] = Pr[A]Pr[B]Pr[C], \ldots, Pr[B \cap C \cap D] = Pr[B]Pr[C]Pr[D]$
- by 4: $Pr[A \cap B \cap C \cap D] = Pr[A]Pr[B]Pr[C]Pr[D]$.

Independence: Question 1

Consider the uniform probability space and the events A, B, C, D.



Which maximal collections of events among A, B, C, D are pairwise independent?

$\{A, B, C\}$, and $\{B, C, D\}$

Can you find three events among A, B, C, D that are mutually independent?

No: We would need an outcome with probability $1/8$.

Independence: Question 2

Let $\Omega = \{1, 2, \ldots, p\}$ be a uniform probability space where p is prime.

Can you find two independent events A and B with $Pr[A], Pr[B] \in (0,1)$?

Let $a = |A|, b = |B|, c = |A \cap B|$.

Then,

$$\begin{aligned} Pr[A \cap B] &= Pr[A]Pr[B], \text{ so that} \\ \frac{c}{p} &= \frac{a}{p} \times \frac{b}{p}. \text{ Hence,} \\ ab &= cp. \end{aligned}$$

This is not possible since $a, b < p$.

Collisions & Collecting

Collisions:

$$Pr[\text{no collision}] \approx e^{-m^2/2n}$$

Collecting:

$$Pr[\text{miss Wilson}] \approx e^{-m/n}$$

$$Pr[\text{miss at least one}] \leq ne^{-m/n}$$

Math Tricks

Approximations:

$$\ln(1-\varepsilon) \approx -\varepsilon$$
$$\exp\{-\varepsilon\} \approx 1-\varepsilon$$

Sums:

$$(a+b)^n = \sum_{m=0}^{n} \binom{n}{m} a^m b^{n-m}$$
$$1 + 2 + \cdots + n = \frac{n(n+1)}{2};$$

Math Tricks, continued

Symmetry: E.g., if we pick balls from a bag, with no replacement,

$$Pr[\text{ball 5 is red}] = Pr[\text{ball 1 is red}]$$

Order of balls = permutation.

All permutations have same probability.

Union Bound:

$$Pr[A \cup B \cup C] \leq Pr[A] + Pr[B] + Pr[C]$$

Inclusion/Exclusion:

$$Pr[A \cup B] = Pr[A] + Pr[B] - Pr[A \cap B]$$

Total Probability:

$$Pr[B] = Pr[A_1]Pr[B|A_1] + \cdots + Pr[A_n]Pr[B|A_n]$$

An L^2-bounded martingale converges almost surely. Just kidding!

A mini-quizz

True or False:

- $Pr[A \cup B] = Pr[A] + Pr[B]$. False True iff disjoint.
- $Pr[A \cap B] = Pr[A]Pr[B]$. False True iff independent.
- $A \cap B = \emptyset \Rightarrow A, B$ independent. False
- For all A, B, one has $Pr[A|B] \geq Pr[A]$. False
- $Pr[A \cap B \cap C] = Pr[A]Pr[B|A]Pr[C|B]$. False

A mini-quizz; part 2

- $\Omega = \{1,2,3,4\}$, uniform. Find events A, B, C that are pairwise independent, not mutually.

  $A = \{1,2\}, B = \{1,3\}, C = \{1,4\}$.
- A, B, C pairwise independent. Is it true that $(A \cap B)$ and C are independent?

  No. In example above, $Pr[A \cap B \cap C] \neq Pr[A \cap B]Pr[C]$.
- Assume $Pr[C|A] > Pr[C|B]$.

  Is it true that $Pr[A|C] > Pr[B|C]$?

  No.
- Deal two cards from a 52-card deck. What is the probability that the value of the first card is strictly larger than that of the second?

  $Pr[\text{same}] = \frac{3}{51}$. $Pr[\text{different}] = \frac{48}{51}$.
  $Pr[\text{first} > \text{second}] = \frac{24}{51}$.

Summary

Good clean fun ....

And good time was had by all ....

Enjoy spring break and the midterm.