CS70: Lecture 20.

Distributions; Independent RVs

1. Review: Expectation
2. Distributions
3. Independent RVs

Review: Expectation

- $E[X] := \sum_x x Pr[X = x] = \sum_\omega X(\omega) Pr[\omega]$.
- $E[g(X,Y)] = \sum_{x,y} g(x,y) Pr[X = x, Y = y]$
  $= \sum_\omega g(X(\omega), Y(\omega)) Pr[\omega]$
- $E[aX + bY + c] = aE[X] + bE[Y] + c$.

Uniform Distribution

Roll a six-sided balanced die. Let X be the number of pips (dots). Then X is equally likely to take any of the values $\{1,2,\ldots,6\}$. We say that X is *uniformly distributed* in $\{1,2,\ldots,6\}$.

More generally, we say that X is uniformly distributed in $\{1,2,\ldots,n\}$ if $Pr[X = m] = 1/n$ for $m = 1,2,\ldots,n$.
In that case,

$$E[X] = \sum_{m=1}^{n} m Pr[X = m] = \sum_{m=1}^{n} m \times \frac{1}{n} = \frac{1}{n}\frac{n(n+1)}{2} = \frac{n+1}{2}.$$

Geometric Distribution

Let's flip a coin with $Pr[H] = p$ until we get H.

For instance:

$\omega_1 = H$, or
$\omega_2 = T\ H$, or
$\omega_3 = T\ T\ H$, or
$\omega_n = T\ T\ T\ T\ \cdots\ T\ H$.

Note that $\Omega = \{\omega_n, n = 1,2,\ldots\}$.

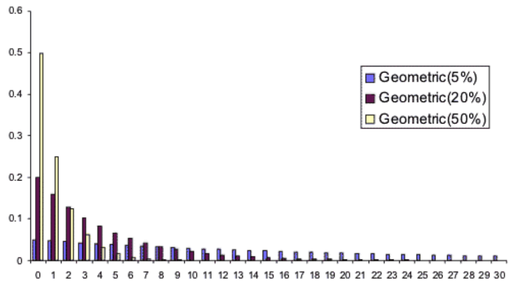
Let X be the number of flips until the first H. Then, $X(\omega_n) = n$.

Also,

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

Geometric Distribution

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



Geometric Distribution

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

Note that

$$\sum_{n=1}^{\infty} Pr[X_n] = \sum_{n=1}^{\infty} (1-p)^{n-1}p = p\sum_{n=1}^{\infty}(1-p)^{n-1} = p\sum_{n=0}^{\infty}(1-p)^n.$$

Now, if $|a| < 1$, then $S := \sum_{n=0}^{\infty} a^n = \frac{1}{1-a}$. Indeed,

$$\begin{aligned} S &= 1 + a + a^2 + a^3 + \cdots \\ aS &= \quad a + a^2 + a^3 + a^4 + \cdots \\ (1-a)S &= 1 + a - a + a^2 - a^2 + \cdots = 1. \end{aligned}$$

Hence,

$$\sum_{n=1}^{\infty} Pr[X_n] = p\frac{1}{1-(1-p)} = 1.$$

Geometric Distribution: Expectation

$$X =_D G(p), \text{ i.e., } Pr[X = n] = (1-p)^{n-1}p, n \geq 1.$$

One has

$$E[X] = \sum_{n=1}^{\infty} nPr[X = n] = \sum_{n=1}^{\infty} n(1-p)^{n-1}p.$$

Thus,

$$\begin{aligned} E[X] &= p + 2(1-p)p + 3(1-p)^2p + 4(1-p)^3p + \cdots \\ (1-p)E[X] &= (1-p)p + 2(1-p)^2p + 3(1-p)^3p + \cdots \\ pE[X] &= p + (1-p)p + (1-p)^2p + (1-p)^3p + \cdots \\ & \quad \text{by subtracting the previous two identities} \\ &= \sum_{n=1}^{\infty} Pr[X = n] = 1. \end{aligned}$$

Hence,

$$E[X] = \frac{1}{p}.$$

Coupon Collectors Problem.

Experiment: Get coupons at random from n until collect all n coupons.
Outcomes: $\{123145..., 56765...\}$
Random Variable: X - length of outcome.

Before: $Pr[X \geq n \ln 2n] \leq \frac{1}{2}$.

Today: $E[X]$?

Time to collect coupons

X-time to get n coupons.

X_1 - time to get first coupon. Note: $X_1 = 1$. $E(X_1) = 1$.

X_2 - time to get second coupon after getting first.

Pr[“get second coupon”|“got ~~milk~~ first coupon”] $= \frac{n-1}{n}$

$E[X_2]$? Geometric ! ! ! $\implies E[X_2] = \frac{1}{p} = \frac{1}{\frac{n-1}{n}} = \frac{n}{n-1}$.

Pr[“getting ith coupon|“got $i-1$rst coupons”] $= \frac{n-(i-1)}{n} = \frac{n-i+1}{n}$

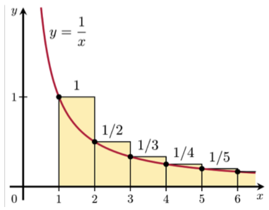
$E[X_i] = \frac{1}{p} = \frac{n}{n-i+1}, i = 1, 2, \ldots, n.$

$$\begin{aligned} E[X] &= E[X_1] + \cdots + E[X_n] = \frac{n}{n} + \frac{n}{n-1} + \frac{n}{n-2} + \cdots + \frac{n}{1} \\ &= n(1 + \frac{1}{2} + \cdots + \frac{1}{n}) =: nH(n) \approx n(\ln n + \gamma) \end{aligned}$$

Review: Harmonic sum

$$H(n) = 1 + \frac{1}{2} + \cdots + \frac{1}{n} \approx \int_1^n \frac{1}{x} dx = \ln(n).$$

.

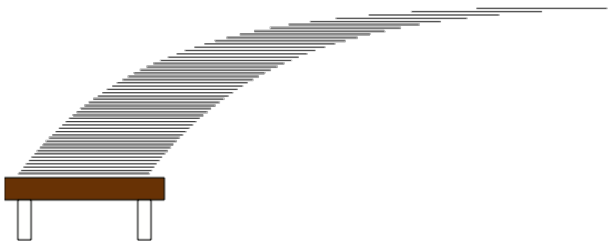


A good approximation is

$H(n) \approx \ln(n) + \gamma$ where $\gamma \approx 0.58$ (Euler-Mascheroni constant).

Harmonic sum: Paradox

Consider this stack of cards (no glue!):



If each card has length 2, the stack can extend $H(n)$ to the right of the table. As n increases, you can go as far as you want!

Paradox

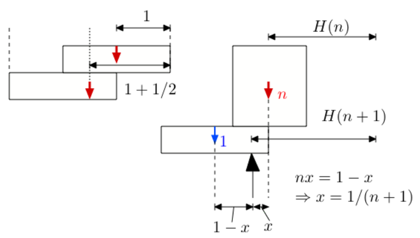
par·a·dox
/ˈperəˌdäks/

noun

a statement or proposition that, despite sound (or apparently sound) reasoning from acceptable premises, leads to a conclusion that seems senseless, logically unacceptable, or self-contradictory.
"a potentially serious conflict between quantum mechanics and the general theory of relativity known as the information paradox"

- a seemingly absurd or self-contradictory statement or proposition that when investigated or explained may prove to be well founded or true.
  "in a paradox, he has discovered that stepping back from his job has increased the rewards he gleans from it"
  synonyms: contradiction, contradiction in terms, self-contradiction, inconsistency, incongruity; More
- a situation, person, or thing that combines contradictory features or qualities.
  "the mingling of deciduous trees with elements of desert flora forms a fascinating ecological paradox"

Stacking



The cards have width 2. Induction shows that the center of gravity after n cards is $H(n)$ away from the right-most edge.

Geometric Distribution: Memoryless

Let X be $G(p)$. Then, for $n \geq 0$,

$$Pr[X > n] = Pr[\text{ first } n \text{ flips are } T] = (1-p)^n.$$

Theorem

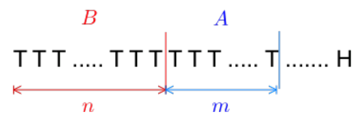
$$Pr[X > n+m | X > n] = Pr[X > m], m, n \geq 0.$$

Proof:

$$\begin{aligned} Pr[X > n+m | X > n] &= \frac{Pr[X > n+m \text{ and } X > n]}{Pr[X > n]} \\ &= \frac{Pr[X > n+m]}{Pr[X > n]} \\ &= \frac{(1-p)^{n+m}}{(1-p)^n} = (1-p)^m \\ &= Pr[X > m]. \end{aligned}$$

Geometric Distribution: Memoryless - Interpretation

$$Pr[X > n+m | X > n] = Pr[X > m], m, n \geq 0.$$



$$Pr[X > n+m | X > n] = Pr[A|B] = Pr[A] = Pr[X > m].$$

The coin is memoryless, therefore, so is X.

Geometric Distribution: Yet another look

Theorem: For a r.v. X that takes the values $\{0, 1, 2, \ldots\}$, one has

$$E[X] = \sum_{i=1}^{\infty} Pr[X \geq i].$$

[See later for a proof.]

If $X = G(p)$, then $Pr[X \geq i] = Pr[X > i-1] = (1-p)^{i-1}$.

Hence,

$$E[X] = \sum_{i=1}^{\infty} (1-p)^{i-1} = \sum_{i=0}^{\infty} (1-p)^i = \frac{1}{1-(1-p)} = \frac{1}{p}.$$

Expected Value of Integer RV

Theorem: For a r.v. X that takes values in $\{0, 1, 2, \ldots\}$, one has

$$E[X] = \sum_{i=1}^{\infty} Pr[X \geq i].$$

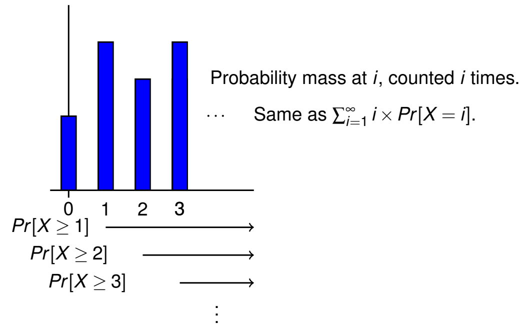
Proof: One has

$$\begin{aligned} E[X] &= \sum_{i=1}^{\infty} i \times Pr[X = i] \\ &= \sum_{i=1}^{\infty} i\{Pr[X \geq i] - Pr[X \geq i+1]\} \\ &= \sum_{i=1}^{\infty} \{i \times Pr[X \geq i] - i \times Pr[X \geq i+1]\} \\ &= \sum_{i=1}^{\infty} \{i \times Pr[X \geq i] - (i-1) \times Pr[X \geq i]\} \\ &= \sum_{i=1}^{\infty} Pr[X \geq i]. \end{aligned}$$

□

Theorem: For a r.v. X that takes values in $\{0, 1, 2, \ldots\}$, one has

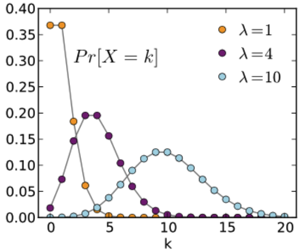
$$E[X] = \sum_{i=1}^{\infty} Pr[X \geq i].$$



Poisson

Experiment: flip a coin n times. The coin is such that $Pr[H] = \lambda/n$.
Random Variable: X - number of heads. Thus, $X = B(n, \lambda/n)$.

Poisson Distribution is distribution of X "for large n."



Poisson

Experiment: flip a coin n times. The coin is such that $Pr[H] = \lambda/n$.
Random Variable: X - number of heads. Thus, $X = B(n, \lambda/n)$.
Poisson Distribution is distribution of X "for large n."
We expect $X \ll n$. For $m \ll n$ one has

$$\begin{aligned} Pr[X=m] &= \binom{n}{m} p^m (1-p)^{n-m}, \text{ with } p = \lambda/n \\ &= \frac{n(n-1)\cdots(n-m+1)}{m!} \left(\frac{\lambda}{n}\right)^m \left(1-\frac{\lambda}{n}\right)^{n-m} \\ &= \frac{n(n-1)\cdots(n-m+1)}{n^m} \frac{\lambda^m}{m!} \left(1-\frac{\lambda}{n}\right)^{n-m} \\ &\approx^{(1)} \frac{\lambda^m}{m!} \left(1-\frac{\lambda}{n}\right)^{n-m} \approx^{(2)} \frac{\lambda^m}{m!} \left(1-\frac{\lambda}{n}\right)^{n} \approx \frac{\lambda^m}{m!} e^{-\lambda}. \end{aligned}$$

For (1) we used $m \ll n$; for (2) we used $(1-a/n)^n \approx e^{-a}$.

Poisson Distribution: Definition and Mean

Definition Poisson Distribution with parameter $\lambda > 0$

$$X = P(\lambda) \Leftrightarrow Pr[X=m] = \frac{\lambda^m}{m!} e^{-\lambda}, m \geq 0.$$

Fact: $E[X] = \lambda$.

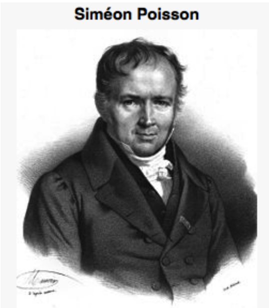
Proof:

$$\begin{aligned} E[X] &= \sum_{m=1}^{\infty} m \times \frac{\lambda^m}{m!} e^{-\lambda} = e^{-\lambda} \sum_{m=1}^{\infty} \frac{\lambda^m}{(m-1)!} \\ &= e^{-\lambda} \sum_{m=0}^{\infty} \frac{\lambda^{m+1}}{m!} = e^{-\lambda}\lambda \sum_{m=0}^{\infty} \frac{\lambda^m}{m!} \\ &= e^{-\lambda}\lambda e^{\lambda} = \lambda. \end{aligned}$$

□

Simeon Poisson

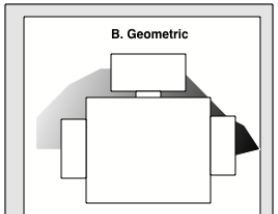
The Poisson distribution is named after:



Siméon Denis Poisson (1781–1840)

Equal Time: B. Geometric

The geometric distribution is named after:



B. Geometric (b. 300 BC)

Prof. Walrand could not find a picture of D. Binomial, sorry.

Review: Distributions

- $U[1,\ldots,n] : Pr[X=m] = \frac{1}{n}, m = 1,\ldots,n;$
  $E[X] = \frac{n+1}{2};$
- $B(n,p) : Pr[X=m] = \binom{n}{m} p^m (1-p)^{n-m}, m = 0,\ldots,n;$
  $E[X] = np;$
- $G(p) : Pr[X=n] = (1-p)^{n-1} p, n = 1,2,\ldots;$
  $E[X] = \frac{1}{p};$
- $P(\lambda) : Pr[X=n] = \frac{\lambda^n}{n!} e^{-\lambda}, n \geq 0;$
  $E[X] = \lambda.$

Independent Random Variables.

Definition: Independence

The random variables X and Y are **independent** if and only if

$$Pr[Y = b|X = a] = Pr[Y = b], \text{ for all } a \text{ and } b.$$

Fact:

X, Y are independent if and only if

$$Pr[X = a, Y = b] = Pr[X = a]Pr[Y = b], \text{ for all } a \text{ and } b.$$

Obvious.

Independence: Examples

Example 1
Roll two die. X, Y = number of pips on the two dice. X, Y are independent.

Indeed: $Pr[X = a, Y = b] = \frac{1}{36}, Pr[X = a] = Pr[Y = b] = \frac{1}{6}$.

Example 2
Roll two die. X = total number of pips, Y = number of pips on die 1 minus number on die 2. X and Y are not independent.

Indeed: $Pr[X = 12, Y = 1] = 0 \neq Pr[X = 12]Pr[Y = 1] > 0$.

Example 3
Flip a fair coin five times, X = number of Hs in first three flips, Y = number of Hs in last two flips. X and Y are independent.

Indeed:

$$Pr[X = a, Y = b] = \binom{3}{a}\binom{2}{b}2^{-5} = \binom{3}{a}2^{-3} \times \binom{2}{b}2^{-2} = Pr[X = a]Pr[Y = b].$$

A useful observation about independence

Theorem

X and Y are independent if and only if

$$Pr[X \in A, Y \in B] = Pr[X \in A]Pr[Y \in B] \text{ for all } A, B \subset \Re.$$

Proof:
If ($\Leftarrow$): Choose $A = \{a\}$ and $B = \{b\}$.

This shows that $Pr[X = a, Y = b] = Pr[X = a]Pr[Y = b]$.

Only if ($\Rightarrow$):

$$\begin{aligned} & Pr[X \in A, Y \in B] \\ &= \sum_{a\in A}\sum_{b\in B} Pr[X = a, Y = b] = \sum_{a\in A}\sum_{b\in B} Pr[X = a]Pr[Y = b] \\ &= \sum_{a\in A}[\sum_{b\in B} Pr[X = a]Pr[Y = b]] = \sum_{a\in A} Pr[X = a][\sum_{b\in B} Pr[Y = b]] \\ &= \sum_{a\in A} Pr[X = a]Pr[Y \in B] = Pr[X \in A]Pr[Y \in B]. \end{aligned}$$

□

Functions of Independent random Variables

Theorem Functions of independent RVs are independent
Let X, Y be independent RV. Then

$$f(X) \text{ and } g(Y) \text{ are independent, for all } f(\cdot), g(\cdot).$$

Proof:
Recall the definition of inverse image:

$$h(z) \in C \Leftrightarrow z \in h^{-1}(C) := \{z \mid h(z) \in C\}. \quad (1)$$

Now,

$$\begin{aligned} & Pr[f(X) \in A, g(Y) \in B] \\ &= Pr[X \in f^{-1}(A), Y \in g^{-1}(B)], \text{ by (1)} \\ &= Pr[X \in f^{-1}(A)]Pr[Y \in g^{-1}(B)], \text{ since } X, Y \text{ ind.} \\ &= Pr[f(X) \in A]Pr[g(Y) \in B], \text{ by (1)}. \end{aligned}$$

□

Mean of product of independent RV

Theorem
Let X, Y be independent RVs. Then

$$E[XY] = E[X]E[Y].$$

Proof:
Recall that $E[g(X, Y)] = \sum_{x,y} g(x, y)Pr[X = x, Y = y]$. Hence,

$$\begin{aligned} E[XY] &= \sum_{x,y} xyPr[X = x, Y = y] = \sum_{x,y} xyPr[X = x]Pr[Y = y], \text{ by ind.} \\ &= \sum_x[\sum_y xyPr[X = x]Pr[Y = y]] = \sum_x[xPr[X = x](\sum_y yPr[Y = y])] \\ &= \sum_x[xPr[X = x]E[Y]] = E[X]E[Y]. \end{aligned}$$

□

Examples

(1) Assume that X, Y, Z are (pairwise) independent, with $E[X] = E[Y] = E[Z] = 0$ and $E[X^2] = E[Y^2] = E[Z^2] = 1$.

Then

$$\begin{aligned} & E[(X + 2Y + 3Z)^2] = E[X^2 + 4Y^2 + 9Z^2 + 4XY + 12YZ + 6XZ] \\ &= 1 + 4 + 9 + 4 \times 0 + 12 \times 0 + 6 \times 0 \\ &= 14. \end{aligned}$$

(2) Let X, Y be independent and $U[1, 2, \ldots n]$. Then

$$\begin{aligned} E[(X - Y)^2] &= E[X^2 + Y^2 - 2XY] = 2E[X^2] - 2E[X]^2 \\ &= \frac{1 + 3n + 2n^2}{3} - \frac{(n+1)^2}{2}. \end{aligned}$$

Mutually Independent Random Variables

Definition

X,Y,Z are mutually independent if

$$Pr[X=x,Y=y,Z=z]=Pr[X=x]Pr[Y=y]Pr[Z=z], \text{ for all } x,y,z.$$

Theorem
The events $A,B,C,\dots$ are pairwise (resp. mutually) independent iff the random variables $1_A,1_B,1_C,\dots$ are pairwise (resp. mutually) independent.
Proof:

$$Pr[1_A=1,1_B=1,1_C=1]=Pr[A\cap B\cap C],\dots$$

□

Functions of pairwise independent RVs

If X,Y,Z are pairwise independent, but not mutually independent, it may be that

$f(X)$ and $g(Y,Z)$ are not independent.

Example 1: Flip two fair coins,
$X=1\{\text{coin 1 is } H\}, Y=1\{\text{coin 2 is } H\}, Z=X\oplus Y$. Then, X,Y,Z are pairwise independent. Let $g(Y,Z)=Y\oplus Z$. Then $g(Y,Z)=X$ is not independent of X.

Example 2: Let A,B,C be pairwise but not mutually independent in a way that A and $B\cap C$ are not independent. Let
$X=1_A, Y=1_B, Z=1_C$. Choose $f(X)=X, g(Y,Z)=YZ$.

Functions of mutually independent RVs

One has the following result:
Theorem
Functions of disjoint collections of mutually independent random variables are mutually independent.
Example:
Let $\{X_n, n\geq 1\}$ be mutually independent. Then,
$Y_1:=X_1X_2(X_3+X_4)^2, Y_2:=\max\{X_5,X_6\}-\min\{X_7,X_8\}, Y_3:=X_9\cos(X_{10}+X_{11})$
are mutually independent.
Proof:
Let $B_1:=\{(x_1,x_2,x_3,x_4)\mid x_1x_2(x_3+x_4)^2\in A_1\}$. Similarly for B_2,B_2.
Then

$$\begin{aligned}&Pr[Y_1\in A_1,Y_2\in A_2,Y_3\in A_3]\\ &=Pr[(X_1,\dots,X_4)\in B_1,(X_5,\dots,X_8)\in B_2,(X_9,\dots,X_{11})\in B_3]\\ &=Pr[(X_1,\dots,X_4)\in B_1]Pr[(X_5,\dots,X_8)\in B_2]Pr[(X_9,\dots,X_{11})\in B_3]\\ &=Pr[Y_1\in A_1]Pr[Y_2\in A_2]Pr[Y_3\in A_3]\end{aligned}$$

□

Operations on Mutually Independent Events

Theorem

Operations on disjoint collections of mutually independent events produce mutually independent events.

For instance, if A,B,C,D,E are mutually independent, then $A\Delta B, C\setminus D, \bar{E}$ are mutually independent.

Proof:

$1_{A\Delta B}=f(1_A,1_B)$ where
$f(0,0)=0, f(1,0)=1, f(0,1)=1, f(1,1)=0$

$1_{C\setminus D}=g(1_C,1_D)$ where
$g(0,0)=0, g(1,0)=1, g(0,1)=0, g(1,1)=0$

$1_{\bar{E}}=h(1_E)$ where
$h(0)=1$ and $h(1)=0$.

Hence, $1_{A\Delta B}, 1_{C\setminus D}, 1_{\bar{E}}$ are functions of mutually independent RVs. Thus, those RVs are mutually independent. Consequently, the events of which they are indicators are mutually independent. □

Product of mutually independent RVs

Theorem
Let $X_1,\dots,X_n$ be mutually independent RVs. Then,

$$E[X_1X_2\cdots X_n]=E[X_1]E[X_2]\cdots E[X_n].$$

Proof:

Assume that the result is true for n. (It is true for $n=2$.)

Then, with $Y=X_1\cdots X_n$, one has

$$\begin{aligned}E[X_1\cdots X_nX_{n+1}] &= E[YX_{n+1}],\\ &= E[Y]E[X_{n+1}],\\ &\quad \text{because } Y, X_{n+1} \text{ are independent}\\ &= E[X_1]\cdots E[X_n]E[X_{n+1}].\end{aligned}$$

□

Summary.

Distributions; Independence

Distributions:

- $G(p): E[X]=1/p$;
- $B(n,p): E[X]=np$;
- $P(\lambda): E[X]=\lambda$

Independence:

- X,Y independent $\Leftrightarrow Pr[X\in A,Y\in B]=Pr[X\in A]Pr[Y\in B]$
- Then, $f(X),g(Y)$ are independent
  and $E[XY]=E[X]E[Y]$
- Mutual independence ....