Random variables

Samy Tindel

Purdue University

Probability - MA 416

Mostly taken from *A first course in probability*
by S. Ross

Outline

Outline

Introduction

Experiment: tossing 3 coins

Model:

$$S = \{h, t\}^3, \; \mathbf{P}(\{s\}) = \frac{1}{8} \text{ for all } s \in S$$

Result of the experiment: we are interested in the quantity

$$X(s) = \text{"\# Heads obtained when } s \text{ is realized"}$$

Introduction (2)

Table for the outcomes:

s	$X(s)$	s	$X(s)$
(t, t, t)	0	(h, t, t)	1
(t, t, h)	1	(h, t, h)	2
(t, h, t)	1	(h, h, t)	2
(t, h, h)	2	(h, h, h)	3

Introduction (3)

Information about X:

X is considered as an application, i.e.

$$X : S \to \{0, 1, 2, 3\}.$$

Then we wish to understand sets like

$$X^{-1}(\{2\}) = \{(t, h, h), (h, t, h), (h, h, t)\}$$

or quantities like

$$\mathbf{P}\left(X^{-1}(\{2\})\right) = \frac{3}{8}.$$

This will be formalized in this chapter

Example: time of first success (1)

Experiment:

- Coin having probability p of coming up heads
- Independent trials: flipping the coin
- Stopping rule: either H occurs or n flips made

Random variable:

$$X = \# \text{ of times the coin is flipped}$$

State space:

$$X \in \{1, \ldots, n\}$$

Example: time of first success (2)

Probabilities for $j < n$:

$$\mathbf{P}(X = j) = \mathbf{P}\left(\{(t, \ldots, t, h)\}\right) = (1 - p)^{j-1} p$$

Probability for $j = n$:

$$\mathbf{P}(X = n) = \mathbf{P}\left(\{(t, \ldots, t, h); (t, \ldots, t, t)\}\right) = (1 - p)^{n-1}$$

Example: time of first success (3)

Checking the sum of probabilities:

$$
\begin{aligned}
\mathbf{P}\left(\bigcup_{j=1}^{n}\{X=j\}\right) &= \sum_{j=1}^{n}\mathbf{P}\left(\{X=j\}\right) \\
&= p\sum_{j=1}^{n-1}(1-p)^{j-1}+(1-p)^{n} \\
&= 1
\end{aligned}
$$

Cumulative distribution function

Definition 1.

Let

- $\mathbf{P}$ a probability on a sample space S
- $X : S \to \mathcal{E}$ a random variable, with $\mathcal{E} \subset \mathbb{R}$

For $x \in \mathbb{R}$ we define

$$F(x) = \mathbf{P}(X \leq x)$$

Then the function F is called cumulative distribution function or distribution function

Outline

General definition

Definition 2.

Let

- $\mathbf{P}$ a probability on a sample space S
- $X : S \to \mathcal{E}$ a random variable

Hypothesis: $\mathcal{E}$ is countable, i.e

$$\mathcal{E} = \{x_i; i \geq 1\}$$

Then we say that X is a discrete random variable

Probability mass function

Definition 3.

Let

- $\mathbf{P}$ a probability on a sample space S
- $\mathcal{E} = \{x_i; i \geq 1\}$ countable state space
- $X : S \to \mathcal{E}$ discrete random variable

For $i \geq 1$ we set

$$p(x_i) = \mathbf{P}(X = x_i)$$

Then the probability mass function of X is the family

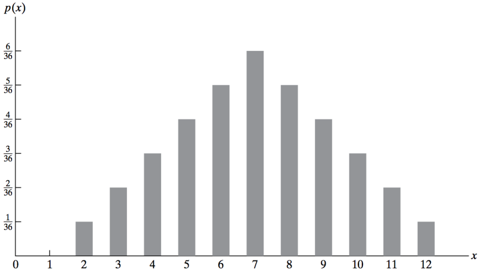
$$\{p(x_i);\ i \geq 1\}$$

Remarks

Sum of the pmf: If p is the pmf of X, then

$$\sum_{i \geq 1} p(x_i) = 1$$

Graph of a pmf: Bar graphs are often used.
Below an example for X = sum of two dice



Example of pmf computation (1)

Definition of the pmf: Let X be a r.v with pmf given by

$$p(i) = c\,\frac{\lambda^i}{i!}, \qquad i \geq 0,$$

where $c > 0$ is a normalizing constant

Question: Compute

1. $\mathbf{P}(X = 0)$
2. $\mathbf{P}(X > 2)$

Example of pmf computation (2)

Computing c: We must have

$$c \sum_{i=0}^{\infty} \frac{\lambda^i}{i!} = 1$$

Thus

$$c = e^{-\lambda}$$

Computing $\mathbf{P}(X = 0)$: We have

$$\mathbf{P}(X = 0) = e^{-\lambda} \frac{\lambda^0}{0!} = e^{-\lambda}$$

Example of pmf computation (3)

Computing $\mathbf{P}(X > 2)$: We have

$$\mathbf{P}(X > 2) = 1 - \mathbf{P}(X \leq 2)$$

Thus

$$\mathbf{P}(X > 2) = 1 - e^{-\lambda}\left(1 + \lambda + \frac{\lambda^2}{2}\right)$$

Cdf for discrete random variables

Proposition 4.

Let

- $\mathbf{P}$ a probability on a sample space S
- $\mathcal{E} = \{x_i; i \geq 1\}$ countable state space, with $\mathcal{E} \subset \mathbb{R}$
- $X : S \to \mathcal{E}$ discrete random variable
- F cdf of X and p pmf of X

Then

1. F can be expressed as

$$F(a) = \sum_{i \geq 1;\, x_i \leq a} p(x_i)$$

2. F is a step function

Example of discrete cdf (1)

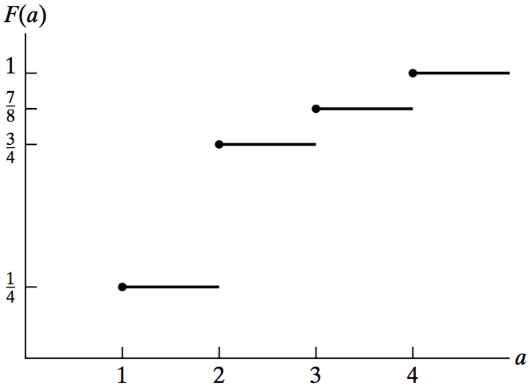
Definition of the random variable:

Consider $X : S \to \{1, 2, 3, 4\}$ given by

$$p(1) = \frac{1}{4}, \quad p(2) = \frac{1}{2}, \quad p(3) = \frac{1}{8}, \quad p(4) = \frac{1}{8}$$

Example of discrete cdf (2)

Graph of F:



Outline

Expected value for discrete random variables

Definition 5.

Let

- $\mathbf{P}$ a probability on a sample space S
- $\mathcal{E} = \{x_i; i \geq 1\}$ countable state space, with $\mathcal{E} \subset \mathbb{R}$
- $X : S \to \mathcal{E}$ discrete random variable
- p pmf of X

Then we define

$$\mathbf{E}[X] = \sum_{i \geq 1} x_i \, \mathbf{P}(X = x_i) = \sum_{i \geq 1} x_i \, p(x_i)$$

Justification of the definition

Experiment:

- Run independent copies of the random variable X
- For i-th copy, the measurement is z_i

Result (to be proved much later):

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} z_i = \mathbf{E}[X]$$

Example: dice rolling (1)

Definition of the random variable: we consider

$$X = \text{outcome when we roll a fair dice}$$

Example: dice rolling (2)

Recall: we consider

$$X = \text{outcome when we roll a fair dice}$$

Pmf: We have $\mathcal{E} = \{1, \ldots, 6\}$ and

$$p(1) = \cdots = p(6) = \frac{1}{6}$$

Expected value: We get

$$\mathbf{E}[X] = \sum_{i=1}^{6} i\, p(i) = \frac{1}{6} \sum_{i=1}^{6} i = \frac{7}{2}$$

Example: indicator of an event (1)

Definition of the random variable:

Let A event with $\mathbf{P}(A) = p$ and set

$$\mathbf{1}_A = \begin{cases} 1 & \text{if } A \text{ occurs} \\ 0 & \text{if } A^c \text{ occurs} \end{cases}$$

Example: indicator of an event (2)

Recall:

Let A event with $\mathbf{P}(A) = p$ and set

$$\mathbf{1}_A = \begin{cases} 1 & \text{if } A \text{ occurs} \\ 0 & \text{if } A^c \text{ occurs} \end{cases}$$

Pmf:

$$p(0) = 1 - p, \qquad p(1) = p$$

Expected value:

$$\mathbf{E}[\mathbf{1}_A] = p$$

Outline

First attempt of a definition

Problem: Let

- X discrete random variable
- $Y = g(X)$ for a function g

How can we compute $\mathbf{E}[g(X)]$?

First strategy:

- $Y = g(X)$ is a discrete random variable
- Determine the pmf p_Y of Y
- Compute $\mathbf{E}[Y]$ according to Definition 5

First attempt: example (1)

Definition of a random variable X:
Let $X : S \to \{-1, 0, 1\}$ with

$$\mathbf{P}(X = -1) = .2, \quad \mathbf{P}(X = 0) = .5, \quad \mathbf{P}(X = 1) = .3$$

We wish to compute $\mathbf{E}[X^2]$

First attempt: example (2)

Definition of a random variable Y: Set $Y = X^2$.
Then $Y \in \{0, 1\}$ and

$$\begin{aligned}\mathbf{P}(Y=0) &= \mathbf{P}(X=0) = .5 \\ \mathbf{P}(Y=1) &= \mathbf{P}(X=-1) + \mathbf{P}(X=1) = .5\end{aligned}$$

First attempt: example (3)

Recall: For $Y = X^2$ we have

$$\mathbf{P}(Y = 0) = .5, \qquad \mathbf{P}(Y = 1) = .5$$

Expected value:

$$\mathbf{E}\left[X^2\right] = \mathbf{E}[Y] = .5$$

Definition of $\mathbf{E}[g(X)]$

Proposition 6.

Let

- X discrete random variable
- p pmf of X
- g real valued function

Then

$$\mathbf{E}\left[g(X)\right] = \sum_{i \geq 1} g(x_i)\, p(x_i) \qquad (1)$$

Proof

Values of Y: We set $Y = g(X)$ and

$$\{y_j;\, j \geq 1\} = \text{ values of } g(x_i) \text{ for } i \geq 1$$

Expression for the rhs of (1): gather according to y_j

$$\begin{aligned}
\sum_{i\geq 1} g(x_i)\, p(x_i) &= \sum_{j\geq 1} \sum_{i;\, g(x_i)=y_j} y_j\, p(x_i) \\
&= \sum_{j\geq 1} y_j \sum_{i;\, g(x_i)=y_j} p(x_i) \\
&= \sum_{j\geq 1} y_j\, \mathbf{P}\left(g(X) = y_j\right) \\
&= \sum_{j\geq 1} y_j\, \mathbf{P}\left(Y = y_j\right) \\
&= \mathbf{E}\left[g(X)\right]
\end{aligned}$$

Previous example reloaded

Definition of a random variable X:
Let $X : S \to \{-1, 0, 1\}$ with

$$\mathbf{P}(X = -1) = .2, \quad \mathbf{P}(X = 0) = .5, \quad \mathbf{P}(X = 1) = .3$$

We wish to compute $\mathbf{E}[X^2]$

Application of (1):

$$\mathbf{E}\left[X^2\right] = \sum_{i=-1,0,1} i^2 p(x_i) = .5$$

Example: seasonal product (1)

Situation:

- Product sold seasonally
- Profit b for each unit sold
- Loss ℓ for each unit left unsold
- Product has to be stocked in advance
  $\hookrightarrow$ s units stocked

Random variable:

- $X = \#$ units of product ordered
- Pmf p for X

Question:
Find optimal s in order to maximize profits

Example: seasonal product (2)

Some random variables: We set

$$X = \#\text{ units ordered, with pmf } p$$
$$Y_s = \text{profit when } s \text{ units stocked}$$

Expression for Y_s:

$$Y_s = (b\,X - (s - X)\,\ell)\,\mathbf{1}_{(X \le s)} + s\,b\,\mathbf{1}_{(X > s)}$$

Expression for $\mathbf{E}[Y_s]$:

$$\mathbf{E}\left[Y_s\right] = \sum_{i=0}^{s} (b\,i - (s - i)\,\ell)\,p(i) + \sum_{i=s+1}^{\infty} s\,b\,p(i)$$

Example: seasonal product (3)

Simplification for $\mathbf{E}[Y_s]$: We get

$$\mathbf{E}[Y_s] = s\,b + (b+\ell)\sum_{i=0}^{s}(i-s)\,p(i)$$

Growth of $s \mapsto \mathbf{E}[Y_s]$: We have

$$\mathbf{E}[Y_{s+1}] - \mathbf{E}[Y_s] = b - (b+\ell)\sum_{i=0}^{s} p(i)$$

Example: seasonal product (4)

Growth of $s \mapsto \mathbf{E}[Y_s]$ (Ctd): We obtain

$$\mathbf{E}\left[Y_{s+1}\right] - \mathbf{E}\left[Y_s\right] > 0 \quad \Longleftrightarrow \quad \sum_{i=0}^{s} p(i) < \frac{b}{b+\ell} \tag{2}$$

Optimization:

- The lhs of (2) is $\nearrow$
- The rhs of (2) is constant
- Thus there exists a s^* such that

$$\mathbf{E}\left[Y_0\right] < \cdots < \mathbf{E}\left[Y_{s^*-1}\right] < \mathbf{E}\left[Y_{s^*}\right] > \mathbf{E}\left[Y_{s^*+1}\right] > \cdots$$

Conclusion: s^* leads to maximal expected profit

Expectation and linear transformations

Proposition 7.

Let

- X discrete random variable
- p pmf of X
- $a, b \in \mathbb{R}$ constants

Then

$$\mathbf{E}[aX + b] = a\,\mathbf{E}[X] + b$$

Proof

Application of relation (1):

$$\begin{aligned}
\mathbf{E}\left[aX+b\right] &= \sum_{i\geq 1}(a\,x_i+b)\;p(x_i) \\
&= a\sum_{i\geq 1}x_i\,p(x_i)+b\sum_{i\geq 1}p(x_i) \\
&= a\,\mathbf{E}\left[X\right]+b
\end{aligned}$$

Outline

Definition of variance

Definition 8.

Let

- X discrete random variable
- p pmf of X
- $\mu = \mathbf{E}[X]$

Then we define $\mathbf{Var}(X)$ by

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

Interpretation

Expected value: For a r.v X, $\mathbf{E}[X]$ represents the mean value of X.

Variance: For a r.v X, $\mathbf{Var}(X)$ represents the dispersion of X wrt its mean value.

A greater $\mathbf{Var}(X)$ means

- The system represented by X has a lot of randomness
- This system is unpredictable

Standard deviation: For physical reasons, it is better to introduce

$$\sigma_X := \sqrt{\mathbf{Var}(X)}.$$

Interpretation (2)

Illustration (from descriptive stats): We wish to compare the performances of 2 soccer players on their last 5 games

Griezmann	5	0	0	0	0
Messi	1	1	1	1	1

Recall: for a set of data $\{x_i;\ i \leq n\}$, we have
Empirical mean: $\bar{x}_n = \frac{1}{n}\sum_{i=1}^{n} x_i$
Empirical variance: $s_n^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x}_n)^2$
Standard deviation: $s_n = \sqrt{s_n^2}$

On our data set: $\bar{x}_G = \bar{x}_M = 1$ goal/game
$\hookrightarrow$ Same goal average
However, $s_G = 2$ goals/game while $s_M = 0$ goals/game
$\hookrightarrow$ M more reliable (less random) than G

Alternative expression for the variance

Proposition 9.

Let

- X discrete random variable
- p pmf of X
- $\mu = \mathbf{E}[X]$

Then $\mathbf{Var}(X)$ can be written as

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

Example: rolling a dice

Random variable:

- X = outcome when one rolls 1 dice

Variance computation: We find

$$\mathbf{E}[X] = \frac{7}{2}, \qquad \mathbf{E}[X^2] = \frac{91}{6}$$

Therefore

$$\mathbf{Var}(X) = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12}$$

Standard deviation:

$$\sigma_X = \sqrt{\frac{35}{12}} \simeq 1.71$$

Variance and linear transformations

Proposition 10.

Let

- X discrete random variable
- p pmf of X
- $a, b \in \mathbb{R}$ constants

Then

$$\mathbf{Var}\,(aX + b) = a^2\,\mathbf{Var}(X)$$

Outline

Bernoulli random variable (1)

Notation:

$$X \sim \mathcal{B}(p) \text{ with } p \in (0,1)$$

State space:

$$\{0,1\}$$

Pmf:

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

Expected value and variance:

$$\mathbf{E}[X] = p, \qquad \mathbf{Var}(X) = p(1-p)$$

Bernoulli random variable (2)

Use 1, success in a binary game:

- Example 1: coin tossing
  - $X = 1$ if H, $X = 0$ if T
  - We get $X \sim \mathcal{B}(1/2)$
- Example 2: dice rolling
  - $X = 1$ if outcome $= 3$, $X = 0$ otherwise
  - We get $X \sim \mathcal{B}(1/6)$

Use 2, answer yes/no in a poll

- $X = 1$ if a person feels optimistic about the future
- $X = 0$ otherwise
- We get $X \sim \mathcal{B}(p)$, with unknown p

Jacob Bernoulli

Some facts about Bernoulli:

- Lifespan: 1654-1705, in Switzerland
- Discovers constant e
- Establishes divergence of $\sum \frac{1}{n}$
- Contributions in diff. eq
- First law of large numbers
- Bernoulli:
  family of 8 prominent mathematicians
- Fierce math fights between brothers

Binomial random variable (1)

Notation:

$$X \sim \text{Bin}(n, p), \text{ for } n \geq 1,\ p \in (0, 1)$$

State space:

$$\{0, 1, \ldots, n\}$$

Pmf:

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

Expected value and variance:

$$\mathbf{E}[X] = np, \qquad \mathbf{Var}(X) = np(1 - p)$$

Binomial random variable (2)

Use 1, Number of successes in a Bernoulli trial:

- Example: Roll a dice 9 times.
- $X = \#$ of 3 obtained
- We get $X \sim \text{Bin}(9, 1/6)$
- $\mathbf{P}(X = 2) = 0.28$

Use 2: Counting a feature in a repeated trial:

- Example: stock of 1000 pants with 10% defects
- Draw 15 times a pant at random
- $X = \#$ of pants with a defect
- We get $X \sim \text{Bin}(15, 1/10)$

Binomial random variable (3)

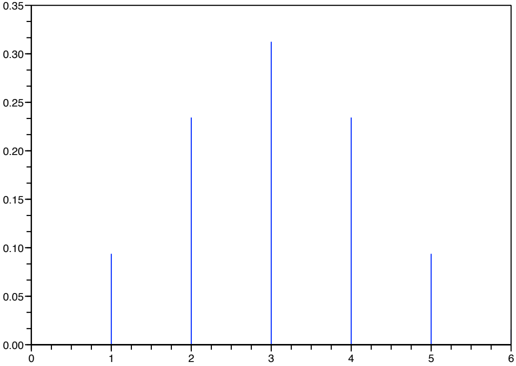


Figure: Pmf for Bin(6; 0.5). x-axis: k. y-axis: $\mathbf{P}(X = k)$

Binomial random variable (4)

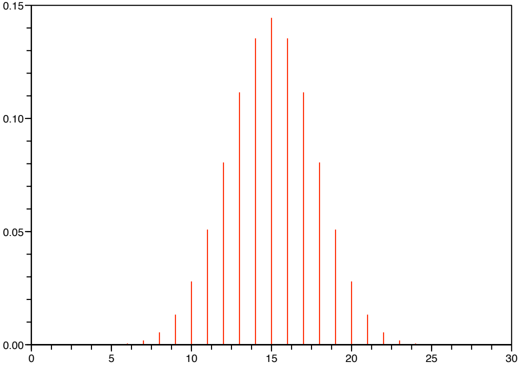


Figure: Pmf for $\text{Bin}(30; 0.5)$. x-axis: k. y-axis: $\mathbf{P}(X = k)$

Example: wheel of fortune (1)

Game:

- Player bets on $1, \ldots, 6$ (say 1)
- 3 dice rolled
- If 1 does not appear, loose \$1
- If 1 appear i times, win \$i

Question:
Find average win

Example: wheel of fortune (2)

Binomial random variable:

- Let $X = \#$ times 1 appears
- Then $X \sim \text{Bin}(3, \frac{1}{6})$

Expression for the win: Set $W =$ win. Then

- $W = \varphi(X)$ with
  $\hookrightarrow \varphi(0) = -1$ and $\varphi(i) = i$ for $i = 1, 2, 3$
- Other expression:

$$W = X - \mathbf{1}_{(X=0)}$$

Example: wheel of fortune (3)

Average win:

$$\mathbf{E}[W] = \mathbf{E}[X] - \mathbf{P}(X=0) = \frac{1}{2} - \left(\frac{5}{6}\right)^3 = -\frac{17}{216}$$

Conclusion: The average win is

$$\mathbf{E}[W] \simeq -\$0.079$$

Pmf variations for a binomial r.v

Proposition 11.

Let

- $X \sim \text{Bin}(n, p)$
- $q = \text{Pmf of } X$
- $k^* = \lfloor (n+1)p \rfloor$

Then we have

- $k \mapsto q(k)$ is $\nearrow$ if $k < k^*$
- $k \mapsto q(k)$ is $\searrow$ if $k > k^*$
- Maximum of q attained for $k = k^*$

Proof

Pmf computation: We have

$$\frac{q(k)}{q(k-1)} = \frac{\mathbf{P}(X=k)}{\mathbf{P}(X=k-1)} = \frac{(n-k+1)p}{k(1-p)}$$

Pmf growth: We get

$$\mathbf{P}(X=k) \geq \mathbf{P}(X=k-1) \quad \Longleftrightarrow \quad k \leq (n+1)p$$

Outline

Poisson random variable (1)

Notation:

$$\mathcal{P}(\lambda) \text{ for } \lambda \in \mathbb{R}_+$$

State space:

$$E = \mathbb{N} \cup \{0\}$$

Pmf:

$$\mathbf{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k \geq 0$$

Expected value and variance:

$$\mathbf{E}[X] = \lambda, \qquad \mathbf{Var}(X) = \lambda$$

Poisson random variable (2)

Use (examples):

- # customers getting into a shop from 2pm to 5pm
- # buses stopping at a bus stop in a period of 35mn
- # jobs reaching a server from 12am to 6am

Empirical rule:
If $n \to \infty$, $p \to 0$ and $np \to \lambda$, we approximate $\text{Bin}(n, p)$ by $\mathcal{P}(\lambda)$.
This is usually applied for

$$p \leq 0.1 \quad \text{and} \quad np \leq 5$$

Poisson random variable (3)

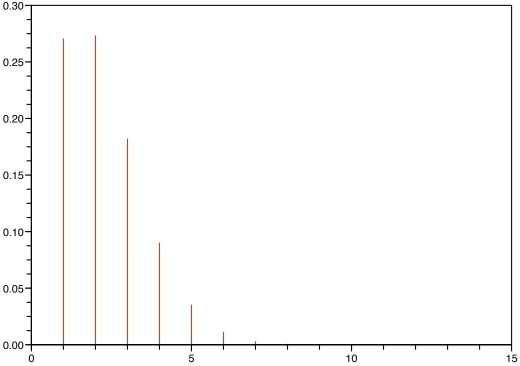


Figure: Pmf of $\mathcal{P}(2)$. x-axis: k. y-axis: $\mathbf{P}(X = k)$

Poisson random variable (4)

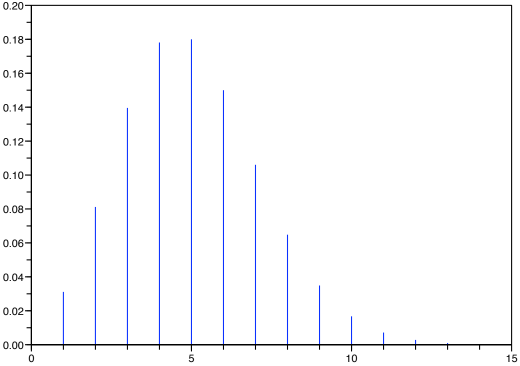
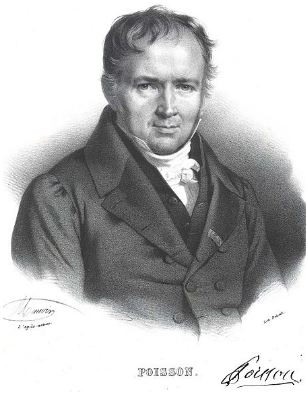


Figure: Pmf of $\mathcal{P}(5)$. x-axis: k. y-axis: $\mathbf{P}(X = k)$

Siméon Poisson

Some facts about Poisson:

- Lifespan: 1781-1840, in $\simeq$ Paris
- Engineer, Physicist and Mathematician
- Breakthroughs in electromagnetism
- Contributions in partial diff. eq
  celestial mechanics, Fourier series
- Marginal contributions in probability



A quote by Poisson:

Life is good for only two things: doing mathematics and teaching it!!

Example: drawing defective items (1)

Experiment:

- Item produced by a certain machine will be defective
  $\hookrightarrow$ with probability .1
- Sample of 10 items drawn

Question:
Probability that the sample contains at most 1 defective item

Example: drawing defective items (2)

Random variable: Let

$$X = \# \text{ of defective items}$$

Then

$$X \sim \text{Bin}(n, p), \quad \text{with} \quad n = 10, \; p = .1$$

Exact probability: We have to compute

$$\begin{aligned} \mathbf{P}(X \leq 1) &= \mathbf{P}(X = 0) + \mathbf{P}(X = 1) \\ &= (0.9)^{10} + 10 \times 0.1 \times (0.9)^{9} \\ &= .7361 \end{aligned}$$

Example: drawing defective items (3)

Approximation: We use

$$\text{Bin}(10, .1) \simeq \mathcal{P}(1)$$

Approximate probability: We have to compute

$$\begin{aligned}
\mathbf{P}(X \leq 1) &= \mathbf{P}(X = 0) + \mathbf{P}(X = 1) \\
&\simeq e^{-1}(1 + 1) \\
&= .7358
\end{aligned}$$

Poisson paradigm

Situation: Consider

- n events $E_1, \ldots, E_n$
- $p_i = \mathbf{P}(E_i)$
- Weak dependence of the E_i: $\mathbf{P}(E_i E_j) \lesssim \frac{1}{n}$
- $\lim_{n\to\infty} \sum_{i=1}^n p_i = \lambda$

Heuristic limit: Under the conditions above we expect that

$$X_n = \sum_{i=1}^n \mathbf{1}_{E_i} \to \mathcal{P}(\lambda) \tag{3}$$

Example: matching problem (1)

Situation:

- n men take off their hats
- Hats are mixed up
- Then each man selects his hat at random
- Match: if a man selects his own hat

Question: Compute

- $\mathbf{P}(E_k)$ with $E_k =$ "exactly k matches"

Example: matching problem (2)

Recall: We have found

$$\mathbf{P}(E_k) = \frac{1}{k!} \sum_{j=2}^{n-k} \frac{(-1)^j}{j!}$$

Thus

$$\lim_{n\to\infty} \mathbf{P}(E_k) = \frac{e^{-1}}{k!}$$

New events: We set

$$G_i = \text{"Person } i \text{ selects his own hat"}$$

Example: matching problem (3)

Probabilities for G_i: We have

$$\mathbf{P}(G_i) = \frac{1}{n}, \qquad \mathbf{P}\left(G_i \middle| G_j\right) = \frac{1}{n-1}$$

Random variable of interest:

$$X = \sum_{i=1}^{n} \mathbf{1}_{G_i} \quad \Longrightarrow \quad \mathbf{P}(E_k) = \mathbf{P}\left(X = k\right)$$

Poisson paradigm: From (3) we have $X \simeq \mathcal{P}(1)$. Therefore

$$\mathbf{P}(E_k) = \mathbf{P}\left(X = k\right) \simeq \mathbf{P}\left(\mathcal{P}(1) = k\right) = \frac{e^{-1}}{k!}$$

Outline

Geometric random variable

Notation:

$$X \sim \mathcal{G}(p), \qquad \text{for } p \in (0,1)$$

State space:

$$E = \mathbb{N} = \{1, 2, 3, \dots\}$$

Pmf:

$$\mathbf{P}(X = k) = p\,(1-p)^{k-1}, \quad k \geq 1$$

Expected value and variance:

$$\mathbf{E}[X] = \frac{1}{p}, \qquad \mathbf{Var}(X) = \frac{1-p}{p^2}$$

Geometric random variable (2)

Use:

- Independent trials, with $\mathbf{P}(\text{success}) = p$
- $X = \#$ trials until first success

Example: dice rolling

- Set $X = $ 1st roll for which outcome $= 6$
- We have $X \sim \mathcal{G}(1/6)$

Computing some probabilities for the example:

$$\begin{aligned} \mathbf{P}(X = 5) &= \left(\frac{5}{6}\right)^4 \frac{1}{6} \simeq 0.08 \\ \mathbf{P}(X \geq 7) &= \left(\frac{5}{6}\right)^6 \simeq 0.33 \end{aligned}$$

Geometric random variable (3)

Computation of $\mathbf{E}[X]$: Set $q = 1 - p$. Then

$$\begin{aligned} \mathbf{E}[X] &= \sum_{i=1}^{\infty} i q^{i-1} p \\ &= \sum_{i=1}^{\infty} (i-1) q^{i-1} p + \sum_{i=1}^{\infty} q^{i-1} p \\ &= q\, \mathbf{E}[X] + 1 \end{aligned}$$

Conclusion:

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

Tail of a geometric random variable

Proposition 12.

Let

- $X \sim \mathcal{G}(p)$
- $n \geq 1$

Then we have

$$\mathbf{P}\left(X \geq n\right) = (1-p)^{n-1}$$

Negative binomial random variable (1)

Notation:

$$X \sim \text{Nbin}(r, p), \text{ for } r \in \mathbb{N}^*, \; p \in (0,1)$$

State space:

$$\{r, r+1, r+2 \ldots\}$$

Pmf:

$$\mathbf{P}(X = k) = \binom{k-1}{r-1} p^r \, (1-p)^{k-r}, \quad k \geq r$$

Expected value and variance:

$$\mathbf{E}[X] = \frac{r}{p}, \qquad \mathbf{Var}(X) = \frac{r(1-p)}{p^2}$$

Negative binomial random variable (2)

Use:

- Independent trials, with $\mathbf{P}(\text{success}) = p$
- $X = \#$ trials until r successes

Justification:

$$(X = k)$$
$$=$$
$$(r - 1 \text{ successes in } (k-1) \text{ 1st trials}) \cap (k\text{-th trial is a success})$$

Thus

$$\mathbf{P}(X = k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}$$

Moments of negative binomial random variable

Proposition 13.

Let

- $X \sim \text{Nbin}(r, p)$, for $r \geq 1$, $p \in (0, 1)$
- $Y \sim \text{Nbin}(r+1, p)$
- $l \geq 1$

Then

$$\mathbf{E}\left[X^l\right] = \frac{r}{p}\,\mathbf{E}\left[(Y-1)^{l-1}\right]$$

Proof (1)

Definition of the l-th moment: We have

$$\mathbf{E}\left[X^l\right] = \sum_{k=r}^{\infty} k^l \binom{k-1}{r-1} p^r (1-p)^{k-r}$$

Relation for combination numbers:

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

Consequence:

$$\mathbf{E}\left[X^l\right] = r\sum_{k=r}^{\infty} k^{l-1} \binom{k}{r} p^r (1-p)^{k-r}$$

Proof (2)

Recall:

$$\mathbf{E}\left[X^l\right] = r\sum_{k=r}^{\infty} k^{l-1}\binom{k}{r} p^r (1-p)^{k-r}$$

From r to $r+1$:

$$\mathbf{E}\left[X^l\right] = \frac{r}{p}\sum_{k=r}^{\infty} k^{l-1}\binom{k}{(r+1)-1} p^{r+1} (1-p)^{(k+1)-(r+1)}$$

Change of variable $j = k+1$:

$$\begin{aligned}\mathbf{E}\left[X^l\right] &= \frac{r}{p}\sum_{j=r+1}^{\infty} (j-1)^{l-1}\binom{j-1}{(r+1)-1} p^{r+1} (1-p)^{j-(r+1)} \\ &= \frac{r}{p}\,\mathbf{E}\left[(Y-1)^{l-1}\right]\end{aligned}$$

Computation of expectation and variance

Consequence of Proposition 13:

$$\mathbf{E}[X] = \frac{r}{p}, \qquad \mathbf{Var}(X) = \frac{r(1-p)}{p^2}$$

The Banach match problem (1)

Situation:

- Pipe smoking mathematician with 2 matchboxes
- 1 box in left hand pocket, 1 box in right hand pocket
- Each time a match is needed, selected at random
- Both boxes contain initially N matches

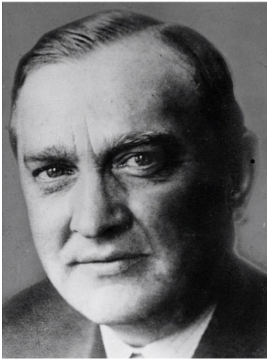
Question:

- When one box is empty,
  what is the probability that k matches are left in the other box?

Stefan Banach

Some facts about Banach:

- Lifespan: 1892-1945, in Krakow and Lviv
- Among greatest 20-th century mathematicians
- Founder of a new field
  $\hookrightarrow$ Functional Analysis
- Survived 2 world wars in tough conditions
- Then dies in 1945 from lung cancer



The Banach match problem (2)

Event: Define E_k by

(Math. discovers that rh box is empty & k matches in lh box)

Expression in terms of a negative binomial:

$$E_k = (X = N + 1 + N - k) = (X = 2N - k + 1),$$

where

$$X \sim \text{Nbin}\left(r = N + 1,\ p = \frac{1}{2}\right)$$

The Banach match problem (3)

Probability of E_k: We get

$$\mathbf{P}(E_k) = \mathbf{P}\left(X = 2N - k + 1\right) = \binom{2N-k}{N}\left(\frac{1}{2}\right)^{2N-k+1}$$

Solution to the problem:
By symmetry between left and right, we get

$$2\,\mathbf{P}(E_k) = \binom{2N-k}{N}\left(\frac{1}{2}\right)^{2N-k}$$

Hypergeometric random variable (1)

Use: Consider the experiment

- Urn containing N balls
- m white balls, $N - m$ black balls
- Sample of size n is drawn without replacement
- Set $X = \#$ white balls drawn

Then

$$X \sim \text{HypG}(n, N, m)$$

Hypergeometric random variable (2)

Notation:

$$X \sim \text{HypG}(n, N, m), \quad \text{for } N \in \mathbb{N}^*,\ m, n \le N,\ p \in (0,1)$$

State space:

$$\{0, \ldots, n\}$$

Pmf:

$$\mathbf{P}(X = k) = \frac{\binom{m}{k}\binom{N-m}{n-k}}{\binom{N}{n}}, \quad 0 \le k \le n$$

Expected value and variance: Set $p = \frac{m}{n}$. Then

$$\mathbf{E}[X] = np, \qquad \mathbf{Var}(X) = np(1-p)\left(1 - \frac{n-1}{N-1}\right)$$

Hypergeometric and binomial

Proposition 14.

Let

- $X \sim \text{HypG}(n, N, m)$,
- Recall that $p = \frac{m}{N}$

Hypothesis:

$$n \ll m, N, \qquad i \ll m, N$$

Then

$$\mathbf{P}(X = i) \simeq \binom{n}{i} p^i (1-p)^{n-i}$$

Proof

Expression for $\mathbf{P}(X = i)$:

$$
\begin{aligned}
\mathbf{P}(X = i) &= \frac{\binom{m}{i}\binom{N-m}{n-i}}{\binom{N}{n}} \\
&= \frac{m!}{(m-i)!i!} \frac{(N-m)!}{(N-m-n+i)!(n-i)!} \frac{(N-n)!n!}{N!} \\
&= \binom{n}{i} \prod_{j=0}^{i-1} \frac{m-j}{N-j} \prod_{k=0}^{n-i-1} \frac{N-m-k}{N-i-k}
\end{aligned}
$$

Approximation: If $i, j, k \ll m, N$ above, we get

$$
\mathbf{P}(X = i) \simeq \binom{n}{i} p^i (1-p)^{n-i}
$$

Example: electric components (1)

Situation: We have

- Lots of electric components of size 10
- We inspect 3 components per lot
  $\hookrightarrow$ Acceptance if all 3 components are non defective
- 30% of lots have 4 defective components
- 70% of lots have 1 defective component

Question:
What is the proportion of rejected lots?

Example: electric components (2)

Events: We define

- A = Acceptance of a lot
- L_1 = Lot with 1 defective component drawn
- L_4 = Lot with 4 defective components drawn

Conditioning: We have

$$\mathbf{P}(A) = \mathbf{P}(A|\, L_1)\, \mathbf{P}(L_1) + \mathbf{P}(A|\, L_4)\, \mathbf{P}(L_4)$$

and

$$\mathbf{P}(L_1) = .7, \qquad \mathbf{P}(L_4) = .3,$$

Example: electric components (3)

Hypergeometric random variable: We check that

$$\mathbf{P}(A|\, L_1) = \mathbf{P}(X_1 = 0), \quad \text{where} \quad X_1 \sim \text{HypG}(3, 10, 1)$$

Thus

$$\mathbf{P}(A|\, L_1) = \frac{\binom{1}{0}\binom{9}{3}}{\binom{10}{3}}$$

Conclusion:

$$\mathbf{P}(A) = \frac{\binom{1}{0}\binom{9}{3}}{\binom{10}{3}} \times 0.7 + \frac{\binom{4}{0}\binom{6}{3}}{\binom{10}{3}} \times 0.3 = 54\%$$

Outline

Another expression for $\mathbf{E}[X]$

Proposition 15.

Let

- $\mathbf{P}$ a probability on a sample space S
- $X : S \to \mathcal{E}$ a random variable

Hypothesis: S is countable, i.e

$$S = \{s_i; i \geq 1\}$$

Then setting $p(s_i) = \mathbf{P}(\{s_i\})$ we have

$$\mathbf{E}[X] = \sum_{i \geq 1} X(s_i) p(s_i)$$

Proof (1)

Recall: We have

$$\mathbf{E}[X] = \sum_{i \geq 1} x_i \, \mathbf{P}(X = x_i)$$

Level set: We define

$$S_i = \{s \in S;\, X(s) = x_i\}$$

Expression for $\mathbf{E}[X]$:

$$\begin{aligned} \mathbf{E}[X] &= \sum_{i \geq 1} x_i \sum_{s_j \in S_i} p(s_j) \\ &= \sum_{i \geq 1} \sum_{s_j \in S_i} X(s_j)\, p(s_j) \end{aligned}$$

Proof (2)

Conclusion: Since $\{S_i;\ i \geq 1\}$ is a partition of S,

$$\mathbf{E}[X] = \sum_{i \geq 1} X(s_i) p(s_i)$$

Expectation of sums

Proposition 16.

Let

- $\mathbf{P}$ a probability on a sample space S
- $X_1, \ldots, X_n : S \to \mathbb{R}$ n random variables

Hypothesis: S is countable, i.e

$$S = \{s_i;\, i \geq 1\}$$

Then

$$\mathbf{E}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathbf{E}\left[X_i\right]$$

Proof

Notation: Set

$$Z = \sum_{i=1}^{n} X_i$$

Expression for $\mathbf{E}[Z]$: According to Proposition 15,

$$\begin{aligned}
\mathbf{E}[Z] &= \sum_{s \in S} Z(s) p(s) \\
&= \sum_{s \in S} \left(\sum_{i=1}^{n} X_i(s) \right) p(s) \\
&= \sum_{i=1}^{n} \left(\sum_{s \in S} X_i(s) p(s) \right) \\
&= \sum_{i=1}^{n} \mathbf{E}\left[X_i\right]
\end{aligned}$$

Example: number of successes (1)

Experiment:

- n trials
- Success for i-th trial with probability p_i
- $X = \#$ of successes

Question:

Expression for $\mathbf{E}[X]$ and $\mathbf{Var}(X)$

Example: number of successes (2)

Expression for X: Let

$$X_i = \mathbf{1}_{(\text{success for } i\text{-th trial})}$$

Then

$$X = \sum_{i=1}^{n} X_i$$

Expression for $\mathbf{E}[X]$: Thanks to Proposition 16, we have

$$\mathbf{E}[X] = \sum_{i=1}^{n} p_i$$

Example: number of successes (3)

Expression for $\mathbf{E}[X^2]$: We invoke the two facts

1. $X_i^2 = X_i$
2. If $i \neq j$, $X_i X_j = \mathbf{1}_{(X_i=1, X_j=1)}$

Therefore

$$\mathbf{E}[X^2] = \sum_{i=1}^{n} \mathbf{E}[X_i^2] + \sum_{i \neq j} \mathbf{E}[X_i X_j]$$

yields

$$\mathbf{E}[X^2] = \sum_{i=1}^{n} p_i + \sum_{i \neq j} \mathbf{P}\left(X_i = 1, X_j = 1\right)$$

Example: number of successes (4)

Particular case, binomial: In this case we have

- The X_i's are independent
- $p_i = p$

New expression for $\mathbf{E}[X^2]$:

$$\mathbf{E}[X^2] = np + n(n-1)p^2$$

Expression for $\mathbf{Var}(X)$:

$$\mathbf{Var}(X) = np(1-p)$$

Example: number of successes (5)

Particular case, hypergeometric: We have

$$\begin{aligned} p_i &= \frac{m}{N} \\ \mathbf{P}\left(X_i = 1, X_j = 1\right) &= \mathbf{P}(X_i = 1)\mathbf{P}\left(X_j = 1 \middle| X_i = 1\right) \\ &= \frac{m}{N}\,\frac{m-1}{N-1} \end{aligned}$$

New expression for $\mathbf{E}[X^2]$:

$$\mathbf{E}[X^2] = np + n(n-1)p\,\frac{m-1}{N-1}$$

Expression for $\mathbf{Var}(X)$:

$$\mathbf{Var}(X) = np(1-p)\left(1 - \frac{n-1}{N-1}\right)$$

Outline

Continuity of the cdf

Proposition 17.

Let

- $\mathbf{P}$ a probability on a sample space S
- $X : S \to \mathcal{E}$ a random variable, with $\mathcal{E} \subset \mathbb{R}$
- F the cdf of X, i.e $F(x) = \mathbf{P}(X \leq x)$

Then the function F satisfies

1. F is a nondecreasing function
2. $\lim_{b\to\infty} F(b) = 1$
3. $\lim_{b\to-\infty} F(b) = 0$
4. F is right continuous

Proof of item 1

Inclusion property: Let $a < b$. Then

$$(X \leq a) \subset (X \leq b)$$

Consequence on probabilities:

$$\mathbf{P}(X \leq a) \leq \mathbf{P}(X \leq b)$$

Proof of item 2

Definition of an increasing sequence: Let $b_n \nearrow \infty$ and

$$E_n = (X \leq b_n)$$

Then

$$\lim_{n\to\infty} E_n = (X < \infty)$$

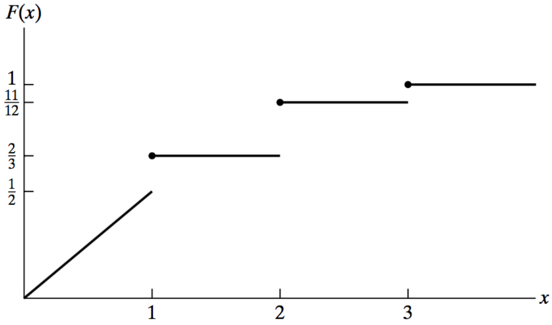
Consequence on probabilities:

$$\begin{aligned} 1 &= \mathbf{P}(X < \infty) \\ &= \mathbf{P}\left(\lim_{n\to\infty} E_n\right) \\ &= \lim_{n\to\infty} \mathbf{P}(E_n) \quad \text{(Since } n \mapsto E_n \text{ is increasing)} \\ &= \lim_{n\to\infty} F(b_n) \end{aligned}$$

Example of cdf (1)

Definition of the function: We set

$$F(x) = \frac{x}{2}\,\mathbf{1}_{[0,1)}(x) + \frac{2}{3}\,\mathbf{1}_{[1,2)}(x) + \frac{11}{12}\,\mathbf{1}_{[2,3)}(x) + \mathbf{1}_{[3,\infty)}(x)$$



Example of cdf (2)

Information read on the cdf: One can check that

- $\mathbf{P}(X < 3) = \frac{11}{12}$
- $\mathbf{P}(X = 1) = \frac{1}{6}$
- $\mathbf{P}(X > \frac{1}{2}) = \frac{3}{4}$
- $\mathbf{P}(2 < X \leq 4) = \frac{1}{12}$