Week 3 Lecture: Discrete random variables

Table of contents

Bayes' Rule

- Simplified version

$$P(F|E) = \frac{P(E \cap F)}{P(E)} = \frac{P(E|F)P(F)}{P(E|F)P(F) + P(E|F^c)P(F^c)}.$$

- General version: $F_1, \ldots, F_n$ is a partition of Ω, event E.

$$P(F_k|E) = \frac{P(E \cap F_k)}{P(E)} = \frac{P(E|F_k)P(F_k)}{\sum_{i=1}^{n} P(E|F_i)P(F_i)}$$

Random Variables

- A **random variable** X on the sample space Ω is a function

$$X : \Omega \to \mathbb{R}.$$

- Roll a die 5 times: $\Rightarrow \Omega = \{(a, b, c, d, e) : a, b, c, d, e \in \{1, \ldots, 6\}\}$
  - X_1 = sum of values which show up

$$X_1((1, 2, 4, 4, 6)) = 1 + 2 + 4 + 4 + 6 = 17.$$

  - X_2 = sum of squares of values which show up

$$X_1((1, 2, 4, 4, 6)) = 1^2 + 2^2 + 4^2 + 4^2 + 6^2 = 73.$$

  - X_3 = number of sixes which show up: $X_3((1, 2, 4, 4, 6)) = 1$.

Discrete Random Variable

- A random variable $X : \Omega \to \mathbb{R}$ is **discrete** if it takes on only countably many values.
- The set of possible values of X, $X(\Omega) = \{X(w) : w \in \Omega\}$, is **countable**. Equivalently, there is an order to list out all elements of this set.

Examples of Discrete Random Variables

- $X =$ number of heads in 3 coin tosses
  $X(\Omega) = \{0, 1, 2, 3\}$ is countable.
- $X =$ number of coin tosses until a head comes up
  $X(\Omega) = \mathbb{Z}^+$ is countable.
- **Remark**: $\mathbb{R}$ is not countable. The proof is beyond the scope of this course. You can *use this property without proof*.

Probability Mass Function (PMF)

- The **probability mass function (PMF)** of a discrete random variable X is a function $p : \mathbb{R} \to [0, 1]$ defined by

$$p(x) = P(X = x).$$

- Capital letters $X, Y, Z, \ldots$ denote random variables. Small letters $x, y, z, \ldots$ denote possible values of $X, Y, Z, \ldots$

Properties of PMF

Lemma 2. Let $X : \Omega \to R$ be a discrete random variable with PMF $p(x)$. Then

(a) $p(x) = 0$ for any $x \notin X(\Omega)$.

(b) $\sum_{x \in X(\Omega)} p(x) = 1$.

Cumulative Distribution Function (CDF)

The **cumulative distribution function (CDF)** of a random variable $X : \Omega \to \mathbb{R}$ is a function $F : \mathbb{R} \to [0, 1]$ defined by

$$F(x) = P(X \leq x), \ x \in \mathbb{R}.$$

Lemma 3. F is a *nondecreasing* function, that is, $F(a) \leq F(b)$ whenever $a \leq b$.

Bernoulli Distribution

- A **Bernoulli trial with success probability** $\mathbf{p}$ is an experiment which has only two outcomes *success* and *failure*

$$P(\text{success}) = p.$$

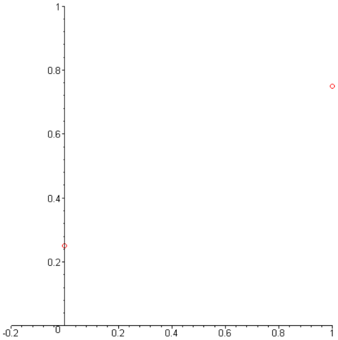
- A **Bernoulli random variable** X counts the number of successes in a Bernoulli trial.
- Write $X \sim \text{Bernoulli}(p)$.

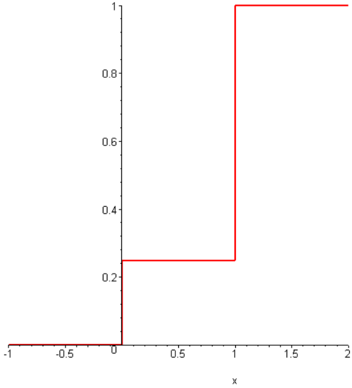
Bernoulli Distribution

- X = number of successes in a Bernoulli trial.
- The set of possible values of X is $\{0, 1\}$.
- It has PMF

$$p(x) = \begin{cases} p \text{ if } x = 1, \\ 1 - p \text{ if } x = 0, \\ 0 \text{ if } x \notin \{0, 1\}. \end{cases}$$

Graph of Bernoulli distribution





PMF and CDF of $X \sim$ Bernoulli$(3/4)$

Binomial Distribution

- A **binomial random variable** X counts the number of successes in n independent Bernoulli trials with success probability p.
- Write $X \sim \text{Binomial}(n, p)$.
- X takes on values $0, 1, \ldots, n$.

PMF of Binomial(n,p)

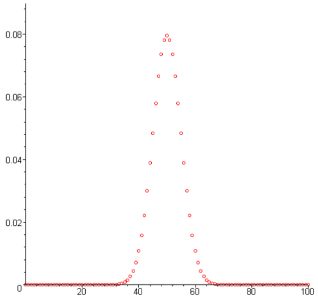
The PMF of $X \sim \text{Binomial}(n, p)$ is

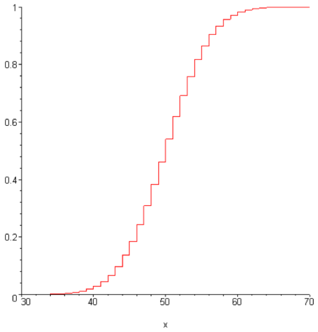
$$p(x) = \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} \text{ if } x \in \{0, 1, \ldots, n\} \\ 0 \text{ otherwise.} \end{cases}$$

Example 1

X = number of heads in 100 independent fair coin tosses

$\Rightarrow X \sim$ Binomial$(100, 0.5)$.





Example 2

One bit (0 or 1) is transmitted 9 times through a noisy channel.

The probability of a failure in a single transmission is 0.05.

The bit is decoded 1 if ≥ 5 ones are received, and 0 otherwise.

What is the probability of incorrect decoding?

Solution.

- Incorrect decoding $\Leftrightarrow$ total number of failures ≥ 5.
- $X =$ total number of failures $\Rightarrow X \sim \text{Binomial}(9, 0.05)$.

$$\begin{aligned} P(\text{incorrect decoding}) &= P(X \geq 5) \\ &= \sum_{x=5}^{9} \binom{9}{x} 0.05^x 0.95^{9-x} \\ &\approx 0.000033. \end{aligned}$$

Example 2 Continued

Geometric Distribution

- Consider a sequence of Bernoulli trials with success prob. p.
- A **geometric random variable** X counts the number of Bernoulli trials needed to get the first success.
- Write $X \sim \text{Geom}(p)$ (or $X \sim \text{Geometric}(p)$).
- The set of possible values for X is $\mathbb{Z}^+$.

PMF of Geom(p)

Lemma 2. $X \sim \text{Geom}(p)$. Then its PMF is

$$p(x) = \begin{cases} (1-p)^{x-1}p \text{ if } x \in \mathbb{Z}^+, \\ 0 \text{ otherwise.} \end{cases}$$

Proof. $X = x$ if the first $x - 1$ trials are failures and the last trial is success.

Example 3

A grandmaster plays a series of chess games against an amateur until the amateur wins a game. The probability that the amateur wins any specific game is 0.001. What is the probability that they are finished after ≤ 100 games?

Example 3 solution

- X = number of games played until the amateur wins the first game $\Rightarrow X \sim \text{Geom}(0.001)$.
- The required probability is

$$P(X \leq 100) = \sum_{x=1}^{100} P(X = x) = \sum_{x=1}^{100} 0.999^{x-1} 0.001 \approx 0.095.$$

Poisson Distribution

- The **Poisson distribution with parameter** λ ($\lambda > 0$) has PMF

$$p(x) = P(X = x) = \frac{\lambda^x}{x!}e^{-\lambda}, \; x = 0, 1, 2, \ldots$$

- Write $X \sim \text{Poisson}(\lambda)$.
- Possible values of X are $0, 1, 2, \ldots$

Rationale of Poisson distribution

- Poisson distribution is used to analyze occurrences in a large number of **very rare events**.
- For example, the total number of car breakdowns among $100,000$ cars in a specific week.

Poisson Distribution vs Binomial Distribution

- Poisson distribution $\approx$ binomial distribution in the case $\mathbf{n}$ **is large** and $\mathbf{p}$ **is small**.
- Recall: $X \sim \text{Binomial}(n, p) \Rightarrow P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}$.
  When n is large and p is small
  - it is hard to compute exactly $\binom{n}{x} p^x (1-p)^{n-x}$,
  - it is difficult to analyze $\text{Binomial}(n, p)$ for large n.

Poisson Distribution Approximates Binomial Distribution

If $X \sim \text{Binomial}(n, p)$ and $\lambda = np$ is a constant, then its PMF is *asymptotically equal* to the PMF of $\text{Poisson}(\lambda)$. Precisely, this means if $\lambda = np$ is a constant, then

$$\lim_{n\to\infty} \binom{n}{x} p^x (1-p)^{n-x} \div \left(\frac{\lambda^x}{x!} e^{-\lambda}\right) = 1.$$

Proof. Optional.

Rule of Thumb

- Lemma 1 says **Binomial**$(\mathbf{n}, \mathbf{p})$ $\approx$**Poisson**$(\mathbf{np})$ when n is large, p is small and np is a constant.
- **Rule of thumb**: The approximation is acceptable if

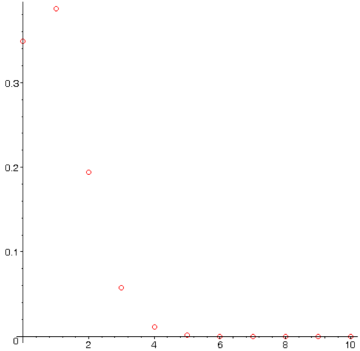
$$\mathbf{n \geq 50 \text{ and } np \leq 10}.$$

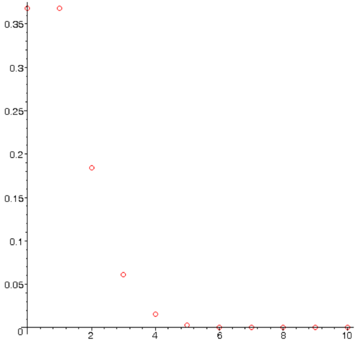
Possion vs Binomial

Poisson distribution is simpler to analyze than Binomial distribution $\Rightarrow$ useful in studying statistical models.

- While Poisson distribution depends only on one parameter λ, binomial distribution depends on two parameters n and p.
- Poisson PMF $p_1(x) = \frac{\lambda^x e^{-\lambda}}{x!}$ is simpler to analyze than binomial PMF $p_2(x) = \binom{n}{x} p^x (1-p)^{n-x}$.

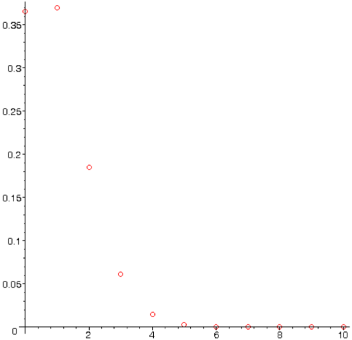
Binomial(10,0.1) vs Poisson(1)

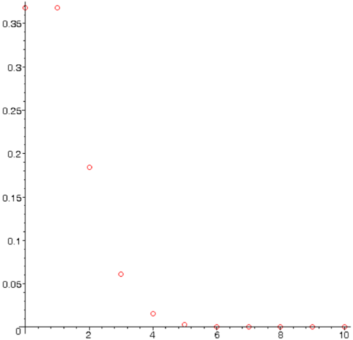




Bad approximation: $n = 10$ is too small

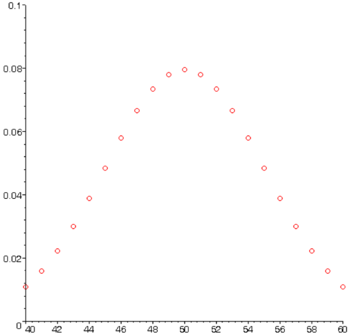
Binomial(100,0.01) vs Poisson(1)

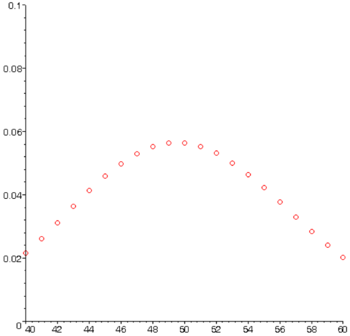




Acceptable approximation: $n = 100 > 50$ and $np = 1 < 10$

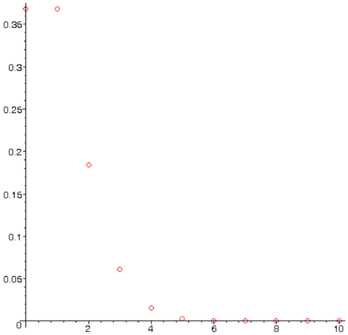
Binomial(100,0.5) vs Poisson(50)

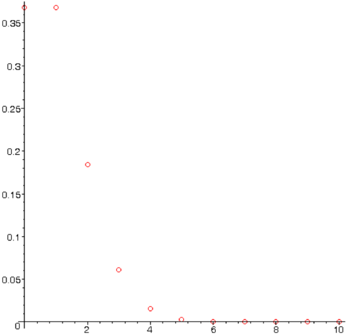




Not really an approximation: $np = 50$ is too large

Binomial(10000,0.0001) vs Poisson(1)





Good approximation: $n = 10,000$ is large and $np = 1$ is small

Basic Integration Formulas

$$\int x^r dx = \frac{x^{r+1}}{r+1} + C, \text{for} \quad r \neq -1$$

$$\int \frac{1}{x} dx = \ln|x| + C, \text{for} \quad x \neq 0$$

$$\int \sin x dx = -\cos x + C \qquad \int \cos x dx = \sin x + C$$

$$\int e^x dx = e^x + C \qquad \int e^{-x} dx = -e^{-x} + C$$

Useful Integration Techniques for Probability

Integration by Parts:

$$\int u dv = uv - \int v du$$

.

To apply integration by parts, you need to make a judicious choice of u and dv so that the integral on the right-hand side is one that you know how to evaluate.

Example

Integration by Parts: Evaluate $\int x \sin x dx$

Let $u = x$ and $dv = \sin x dx$

Example: Poor choice of u and dv

Let $u = \sin x$ and $dv = xdx$

Example

Integration by Parts: Evaluate $\int \ln x dx$

Let $u = \ln x$ and $dv = dx$

Repeated Integration by Parts

Evaluate $\int x^2 \sin x dx$.

Continue

Repeated Integration by Parts with a Twist

Evaluate $\int e^{2x} \sin x dx$
Let $u = e^{2x}$ and $dv = \sin x dx$

Continue

Definite Integrals for Integration by Parts

$$\int_{x=a}^{x=b} u dv = uv|_{x=a}^{x=b} - \int_{x=a}^{x=b} v du$$

Example: Evaluate $\int_1^2 x^3 \ln x dx$

Continue

Improper Integrals

If f is continuous on the interval $[a, \infty)$, we define the improper integral $\int_a^\infty f(x)dx$ to be

$$\int_a^\infty f(x)dx = \lim_{R \to \infty} \int_a^R f(x)dx$$

Quiz 1

- Week 4 during tutorial
- 50 minutes in total
- 10 MCQ
- Weight: 10%
- Contents tested: Week 1-3 Lecture Notes
- Calculators are allowed.
- For Quiz 1 **No** Formula sheet will be given