Discrete Random Variables

Statistics I · Section 5

Zong-Wei Yeh

Department of Finance, National Chung Cheng University

Fall 2026

Outline

5.1 The Bernoulli distribution
5.2 The binomial distribution
5.3 The geometric distribution
5.4 The hypergeometric distribution
5.5 The Poisson distribution
5.6 How the five fit together

The Bernoulli Distribution

The Bernoulli distribution

Definition 5.1 (Bernoulli)

X is *Bernoulli* (伯努利分配), written $X \sim \text{Bernoulli}(p)$, if $X \in \{0, 1\}$ with $p \in [0, 1]$, $P(X = 1) = p$ and $P(X = 0) = q = 1 - p$. Its pmf is $p(x) = p^x q^{1-x}$.

- A Bernoulli variable models a single yes/no trial, such as a default or a defect.
- Most of the other models in this section are built from it.

Bernoulli MGF, mean, variance

Proposition 5.2

If $X \sim \text{Bernoulli}(p)$: $M_X(t) = q + pe^t$; $\mathbb{E}(X) = p$, $\text{Var}(X) = pq$.

- Both come from $M_X'(0)$ and $M_X''(0)$.
- The variance is largest near $p = \frac{1}{2}$.

Example 5.3

Example 5.3

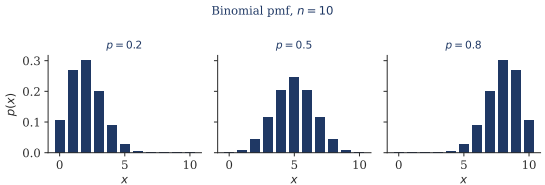
A loan defaults with probability $p = 0.02$; $X = 1$ if it defaults. Write the pmf and find $\mathbb{E}(X)$, $\mathrm{Var}(X)$, σ_X.

The Binomial Distribution

The binomial distribution

Definition 5.4 (Binomial)

X is *binomial* (二項分配) if it counts the successes in n independent Bernoulli(p) trials, so $X \sim \text{Bin}(n, p)$ with $p(x) = \binom{n}{x} p^x q^{n-x}$, $x = 0, \ldots, n$.



- The pmf multiplies two things. The first is the probability of one particular sequence. The second counts how many there are.

Binomial: a valid pmf

Proposition 5.5

$$\sum_{x=0}^{n} \binom{n}{x} p^x q^{n-x} = (p+q)^n = 1.$$

- Immediate from the binomial theorem, since $p + q = 1$.

Binomial MGF

Proposition 5.6

$M_X(t) = (q + pe^t)^n$.

- Straight from the pmf: $\sum_x \binom{n}{x}(pe^t)^x q^{n-x}$ is $(q + pe^t)^n$ by the binomial theorem.
- We do *not* write X as a sum of Bernoullis here. The next result proves that *from* this MGF.

A sum of Bernoullis is binomial

Proposition 5.7

If $X_1, \ldots, X_n$ are independent Bernoulli(p), then $X_1 + \cdots + X_n \sim \text{Bin}(n, p)$.

- Each MGF is $q + pe^t$. Independence multiplies them (Proposition 4.80) to $(q + pe^t)^n$, and uniqueness (Theorem 4.42) names the distribution.
- This is where the binomial pmf comes from. Counting and the MGF are two routes to one answer.

Binomial mean and variance

Proposition 5.8

$\mathbb{E}(X) = np, \quad \mathrm{Var}(X) = npq.$

- Read them off M_X, or get them in one line from the sum of Bernoullis: means always add, and variances add when the terms are independent.

Example 5.9

Example 5.9

$X \sim \text{Bin}(10, 0.3)$. Find $P(X = 3)$, $P(X \leq 2)$, and $\mathbb{E}(X)$, $\text{Var}(X)$, σ_X.

The Geometric Distribution

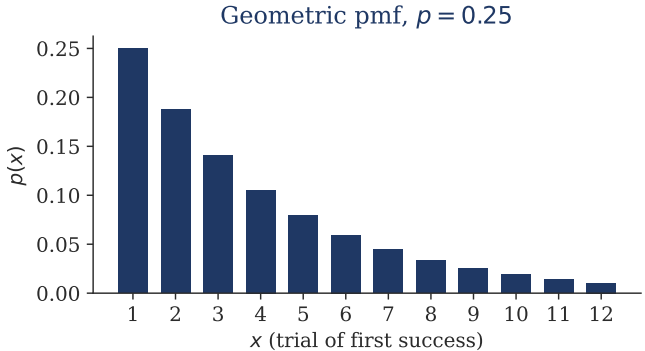
The geometric distribution

Definition 5.10 (Geometric)

X is *geometric* (幾何分配) if it counts the independent Bernoulli(p) trials up to and including the first success, with $p \in (0, 1)$. We write $X \sim \text{Geometric}(p)$. It has support $\{1, 2, \dots\}$ and pmf $p(x) = q^{x-1}p$.

- The pmf multiplies two things. The first is the probability that the early trials all fail. The second is the probability that the last trial succeeds.

The shape of the geometric pmf



- Each bar is a factor q shorter than the one before it.

Geometric: valid pmf and tail

Proposition 5.11

$$\sum_{x=1}^{\infty} q^{x-1}p = 1, \quad \text{and } P(X > x) = q^{x} \quad \text{for every integer } x \geq 0.$$

- A geometric series gives the first.
- For the second, $X > x$ just means the first x trials all fail.

Geometric MGF, mean, variance

Proposition 5.12

$M_X(t) = \dfrac{pe^t}{1 - qe^t}$ for $qe^t < 1$; $\mathbb{E}(X) = 1/p$, $\mathrm{Var}(X) = \dfrac{q}{p^2}$.

- Sum a geometric series in e^t, then differentiate at 0.

A process with no memory

Theorem 5.13 (Memorylessness)

For integers $m, n \geq 0$, $P(X > m + n \mid X > m) = P(X > n)$.

- **Memoryless** (無記憶性): past failures never make a success "due".
- *Alternative convention:* counting failures *before* the first success gives $Y = X - 1$, with mean q/p. This course uses X, with mean $1/p$.

Example 5.14

Example 5.14

Each drilling attempt strikes oil with $p = 0.25$. Let X be the attempt of the first strike. Find $P(X = 3)$, $\mathbb{E}(X)$, and $P(X > 4)$.

The Hypergeometric Distribution

The hypergeometric distribution

Definition 5.15 (Hypergeometric) $X \sim \textbf{Hypergeometric}(N, K, n)$

A population of $N \geq 2$ holds K successes, $0 \leq K \leq N$. Draw n, $1 \leq n \leq N$, without replacement and let X count the successes drawn. Then X is *hypergeometric* (超幾何分配), with

$$p(x) = \frac{\binom{K}{x}\binom{N-K}{n-x}}{\binom{N}{n}}, \quad \max(0, n-(N-K)) \leq x \leq \min(n, K).$$

- Drawing without replacement makes the draws dependent, so the count is no longer binomial.

Hypergeometric mean and variance

Proposition 5.16 **($\pi = K/N$)**

$\mathbb{E}(X) = nK/N$, $\mathrm{Var}(X) = n\,\pi(1-\pi)\,\frac{N-n}{N-1}$.

- Write $X = \sum I_i$ with indicators. Means add, whatever the dependence.
- The variance is stated without proof.
- **The extra factor** in the variance is the finite-population correction (有限母體修正). It is below 1 for $n \geq 2$, because dependent draws spread less.

Example 5.17

Example 5.17

A batch of $N = 50$ parts has $K = 20$ defectives; inspect $n = 5$ without replacement. Find $P(X = 2)$, $\mathbb{E}(X)$, $\mathrm{Var}(X)$.

When the population is large

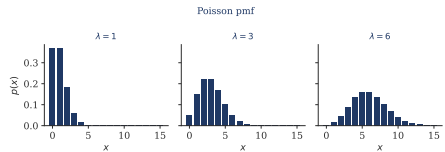
- **Remark (binomial approximation).** If N is large relative to n, removing a few items barely changes the population. The draws are then almost independent. The hypergeometric is then close to $\text{Bin}(n, \pi)$.
- Precisely: if $N, K \to \infty$ with $K/N \to \pi$, and n, x stay fixed, then $p(x) \to \binom{n}{x} \pi^x (1-\pi)^{n-x}$. We state this without proof.
- It improves fast. At $n = 5$, $\pi = 0.2$: $P(X = 1)$ is 0.4313, 0.4148, 0.4097 at $N = 50, 200, 10\,000$, against the binomial's 0.4096.

The Poisson Distribution

The Poisson distribution

Definition 5.18 (Poisson)

X is *Poisson* (卜瓦松分配) with parameter $\lambda > 0$, written $X \sim \text{Poisson}(\lambda)$, if it has support $\{0, 1, 2, \dots\}$ and pmf $p(x) = \frac{e^{-\lambda}\lambda^x}{x!}$.



- **Where it comes from.** Counts of independent events at an average rate (發生率) λ per unit, arriving one at a time.

Poisson: a valid pmf

Proposition 5.19

$$\sum_{x=0}^{\infty} \frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda}e^{\lambda} = 1.$$

- The exponential series, after factoring out $e^{-\lambda}$.

Poisson MGF, mean, variance

Proposition 5.20

$M_X(t) = e^{\lambda(e^t-1)}$, $\mathbb{E}(X) = \mathrm{Var}(X) = \lambda$.

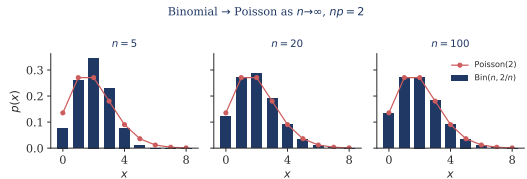
- Sum the exponential series with argument λe^t, then differentiate at 0.

The law of rare events

Theorem 5.21 (Poisson limit of the binomial)

Fix $\lambda > 0$ and let $n \to \infty$ with $p = \lambda/n$. Then for each fixed x,

$$\binom{n}{x} p^x (1-p)^{n-x} \to \frac{e^{-\lambda}\lambda^x}{x!}.$$



- Many trials, each one rare, with the expected count held fixed.

Sum of independent Poissons

Proposition 5.22

If $X \sim \text{Poisson}(\lambda_1)$ and $Y \sim \text{Poisson}(\lambda_2)$ are independent, then $X + Y \sim \text{Poisson}(\lambda_1 + \lambda_2)$.

- The MGFs multiply (Proposition 4.63), and uniqueness names the result.
- Pooling two independent streams simply adds their rates.

Example 5.23

Example 5.23

A help desk receives calls at $\lambda = 3$ per hour (Poisson). Find $P(X = 0)$, $P(X = 2)$, and $P(X \geq 1)$.

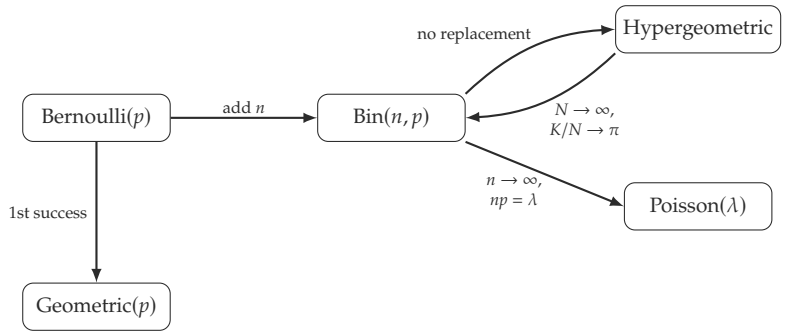
Example 5.24

Example 5.24

$n = 200$ policies each lapse independently with $p = 0.01$. Approximate $P(\text{no lapses})$ by a Poisson and compare with the exact binomial.

How the Five Fit Together

The map



– Down from the Bernoulli changes what we count. Out of the binomial changes how we sample or takes a limit.

Which model?

Ask one question first: **what does X count?**

X **counts**	**Check**	**Model**
successes in n trials	independent, with replacement	Binomial
successes in n draws	without replacement	Hypergeometric
trials until the first success	independent trials	Geometric
events in a window of time	no trials, only a rate	Poisson

- The binomial fixes the trials and counts successes. The geometric fixes one success and counts trials.