Famous Distributions

The Characters You'll Meet Again and Again

Scott Cunningham

Harvard University
Department of Government

February 5, 2026

Where Are We?

Monday: Random variables

- Random variables as functions from outcomes to numbers
- PMFs (discrete), PDFs (continuous), CDFs (both)
- Joint distributions and independence of random variables

Today: Famous distributions

- Discrete: Bernoulli, Binomial, Poisson
- Continuous: Uniform, Normal, Exponential
- Why these specific distributions matter

Reading: Aronow & Miller §1.2 (pp. 15–31), Blackwell Ch. 2.3

Why These Distributions?

Many distributions are named. We focus on six because:

1. **They model real phenomena**: Elections, counts, durations, measurements
2. **They're mathematically tractable**: We can derive expectations, variances
3. **They recur constantly**: Master these, and you're equipped for the course

Today’s Cast of Characters

- **Bernoulli** — the foundation (yes/no outcomes)
- **Binomial** — counting successes
- **Poisson** — rare events
- **Uniform** — equiprobable outcomes
- **Normal** — the star of the show
- **Exponential** — waiting times

The Bernoulli Distribution

Success or Failure

The Bernoulli Distribution

The simplest random variable: **two outcomes**.

Definition: $X \sim \text{Bernoulli}(p)$ if:

$$X = \begin{cases} 1 & \text{with probability } p \\ 0 & \text{with probability } 1 - p \end{cases}$$

The PMF is: $f_X(x) = p^x(1-p)^{1-x}$ for $x \in \{0, 1\}$

Examples:

- Coin flip: $p = 0.5$
- Voter turnout: Did this person vote? ($p \approx 0.6$ in US)
- Survey response: Does respondent approve? ($p =?$)

Bernoulli: Expectation and Variance

For $X \sim \text{Bernoulli}(p)$:

Expected value:

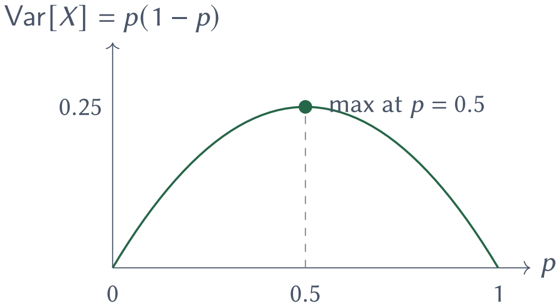
$$\mathbb{E}[X] = 0 \cdot (1 - p) + 1 \cdot p = p$$

Variance:

$$\begin{aligned} \text{Var}[X] &= \mathbb{E}[X^2] - (\mathbb{E}[X])^2 \\ &= p - p^2 = p(1 - p) \end{aligned}$$

Note: Variance is maximized when $p = 0.5$. Certainty ($p = 0$ or $p = 1$) means zero variance.

Visualizing Bernoulli Variance



When outcomes are most uncertain ($p = 0.5$), variance is highest.

The Binomial Distribution

Counting Successes

From Bernoulli to Binomial

Setup: Run n independent Bernoulli trials, each with success probability p.

Question: What's the distribution of the **total number of successes**?

If $X_1, X_2, \ldots, X_n \overset{\text{iid}}{\sim} \text{Bernoulli}(p)$, then:

$$Y = \sum_{i=1}^{n} X_i \sim \text{Binomial}(n, p)$$

Examples:

- 10 coin flips: How many heads?
- 1000 voters sampled: How many support candidate A?
- 50 precincts: How many have irregularities?

The Binomial PMF

For $Y \sim \text{Binomial}(n, p)$:

$$f_Y(k) = \mathbb{P}(Y = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

for $k \in \{0, 1, 2, \ldots, n\}$

Why this formula?

- p^k: probability of k successes
- $(1-p)^{n-k}$: probability of $n-k$ failures
- $\binom{n}{k}$: number of ways to arrange k successes in n trials

The binomial coefficient "chooses" which trials are successes.

Binomial: Expectation and Variance

For $Y \sim \text{Binomial}(n, p)$:

Expected value: Since $Y = \sum_{i=1}^{n} X_i$ where $X_i \sim \text{Bernoulli}(p)$:

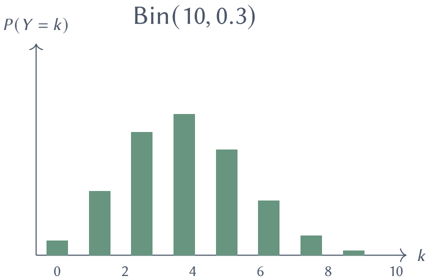
$$\mathbb{E}[Y] = \sum_{i=1}^{n} \mathbb{E}[X_i] = np$$

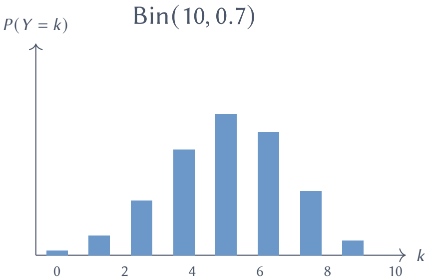
Variance: Since the X_i are independent:

$$\text{Var}[Y] = \sum_{i=1}^{n} \text{Var}[X_i] = np(1-p)$$

Linearity of expectation works always. Additivity of variance requires independence.

Visualizing the Binomial





The distribution is centered at np and symmetric when $p = 0.5$.

Part III

The Poisson Distribution

Counts of Rare Events

The Poisson Distribution

Setup: Counting events that occur **independently** at a constant **rate**.

Definition: $X \sim \text{Poisson}(\lambda)$ has PMF:

$$f_X(k) = \frac{\lambda^k e^{-\lambda}}{k!} \quad \text{for } k \in \{0, 1, 2, \ldots\}$$

where $\lambda > 0$ is the **rate parameter**.

Examples:

- Number of coups in a region per decade
- Number of Supreme Court vacancies per presidential term
- Number of mass casualty events per year

Poisson: Key Properties

For $X \sim \text{Poisson}(\lambda)$:

Expected value: $\mathbb{E}[X] = \lambda$

Variance: $\text{Var}[X] = \lambda$

The mean equals the variance. This is the defining characteristic.

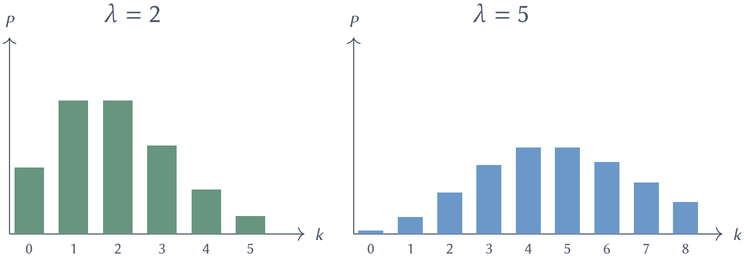
Poisson approximation to Binomial:

If n is large and p is small (so $np = \lambda$ is moderate):

$$\text{Binomial}(n, p) \approx \text{Poisson}(\lambda = np)$$

This is useful for rare events: many trials, low probability per trial.

Visualizing the Poisson



As λ increases, the distribution shifts right and becomes more symmetric.

The Uniform Distribution

All Outcomes Equally Likely

The Continuous Uniform Distribution

Definition: $X \sim \text{Uniform}(a, b)$ has PDF:

$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{if } a \leq x \leq b \\ 0 & \text{otherwise} \end{cases}$$

Every value in $[a, b]$ is equally likely.

CDF:

$$F_X(x) = \begin{cases} 0 & \text{if } x < a \\ \frac{x-a}{b-a} & \text{if } a \leq x \leq b \\ 1 & \text{if } x > b \end{cases}$$

Uniform: Expectation and Variance

For $X \sim \text{Uniform}(a, b)$:

Expected value:

$$\mathbb{E}[X] = \frac{a+b}{2}$$

(The midpoint — by symmetry)

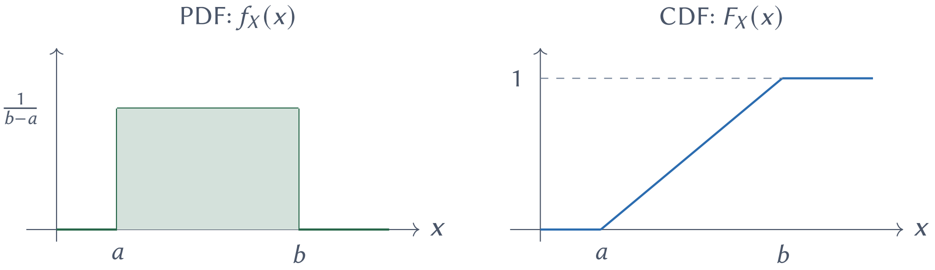
Variance:

$$\text{Var}[X] = \frac{(b-a)^2}{12}$$

Special case: $\text{Uniform}(0, 1)$ has $\mathbb{E}[X] = 0.5$ and $\text{Var}[X] = 1/12$.

Political science example: Random assignment in experiments. If we randomly assign treatment with probability 0.5, we're implicitly drawing from $\text{Uniform}(0, 1)$ and treating if the draw < 0.5.

Visualizing the Uniform



Flat PDF means equal probability density everywhere in $[a, b]$.

Part V

The Normal Distribution

The Star of the Show

The Normal Distribution

Definition: $X \sim \text{Normal}(\mu, \sigma^2)$ has PDF:

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

for $x \in \mathbb{R}$, where μ is the mean and σ^2 is the variance.

Why so important?

1. **Central Limit Theorem**: Sample means are approximately normal
2. **Mathematical convenience**: Closed under addition, scaling
3. **Good approximation**: Economic indicators, polling aggregates, measurement error

Normal: Key Properties

For $X \sim \text{Normal}(\mu, \sigma^2)$:

Expected value: $\mathbb{E}[X] = \mu$

Variance: $\text{Var}[X] = \sigma^2$

Closure properties:

- If $X \sim N(\mu, \sigma^2)$, then $aX + b \sim N(a\mu + b, a^2\sigma^2)$
- If $X \sim N(\mu_X, \sigma_X^2)$ and $Y \sim N(\mu_Y, \sigma_Y^2)$ are independent, then $X + Y \sim N(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2)$

These properties make the normal uniquely tractable for statistical inference.

The Standard Normal

Definition: $Z \sim N(0, 1)$ is the **standard normal**.

Standardization: If $X \sim N(\mu, \sigma^2)$, then:

$$Z = \frac{X - \mu}{\sigma} \sim N(0, 1)$$

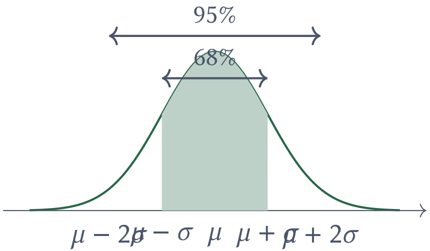
Why standardize?

- Tables and software give probabilities for Z
- Comparing variables on different scales
- Building test statistics

Convention: $\Phi(z) = P(Z \leq z)$ is the standard normal CDF.

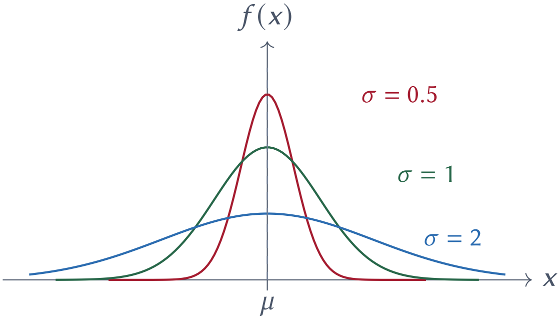
The 68–95–99.7 Rule

For $X \sim N(\mu, \sigma^2)$:



- 68% of values within 1 SD of mean
- 95% within 2 SDs (actually 1.96)
- 99.7% within 3 SDs

Visualizing the Normal



Larger $\sigma \implies$ flatter, more spread out. Same total area (= 1) always.

Part VI

The Exponential Distribution

Waiting for an Event

The Exponential Distribution

Setup: How long until the next event, if events occur at constant rate λ?

Definition: $X \sim \text{Exponential}(\lambda)$ has PDF:

$$f_X(x) = \lambda e^{-\lambda x} \quad \text{for } x \geq 0$$

where $\lambda > 0$ is the **rate parameter** (same λ as Poisson).

Political science examples:

- Time until a cabinet collapse
- Duration of ceasefires
- Time between policy changes

Exponential: Key Properties

For $X \sim \text{Exponential}(\lambda)$:

Expected value: $\mathbb{E}[X] = \frac{1}{\lambda}$

Variance: $\text{Var}[X] = \frac{1}{\lambda^2}$

CDF: $F_X(x) = 1 - e^{-\lambda x}$ for $x \geq 0$

Memoryless property: $\mathbb{P}(X > s + t \mid X > s) = \mathbb{P}(X > t)$

The probability of waiting another t units doesn't depend on how long you've already waited. This is the continuous analog of the geometric distribution.

Poisson and Exponential: Two Sides of One Coin

The Poisson–Exponential connection:

If events occur at rate λ:

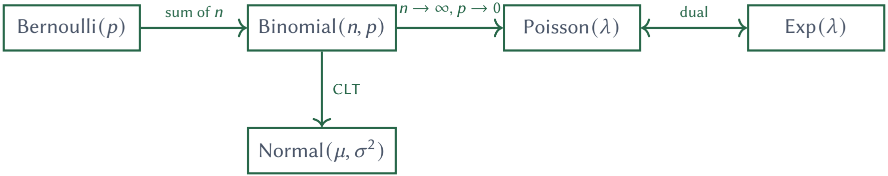
- **Number of events** in time $t \sim \text{Poisson}(\lambda t)$
- **Time between events** $\sim \text{Exponential}(\lambda)$

Same process, different questions.

Example: Supreme Court vacancies

- Poisson: How many vacancies in 4 years?
- Exponential: How long until the next vacancy?

How These Distributions Connect



- Bernoulli $\rightarrow$ Binomial: Sum of independent trials
- Binomial $\rightarrow$ Poisson: Many trials, small probability
- Binomial $\rightarrow$ Normal: Central Limit Theorem (Week 5)
- Poisson $\leftrightarrow$ Exponential: Counts vs. waiting times

Summary: Six Distributions to Know

Distribution	**Support**	$\mathbb{E}[X]$	$\text{Var}[X]$	**Use case**
Bernoulli(p)	$\{0, 1\}$	p	$p(1-p)$	Binary outcomes
Binomial(n, p)	$\{0, \ldots, n\}$	np	$np(1-p)$	Count successes
Poisson(λ)	$\{0, 1, 2, \ldots\}$	λ	λ	Rare event counts
Exponential(λ)	$[0, \infty)$	$1/\lambda$	$1/\lambda^2$	Waiting times
Uniform(a, b)	$[a, b]$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$	Equal probability
Normal(μ, σ^2)	$\mathbb{R}$	μ	σ^2	The default

These distributions are the **vocabulary** of statistics. *Master them now.*

Working with Distributions in R

Simulating and Visualizing

Sampling from Distributions in R

R has functions for every major distribution:

```
# Generate random samples
rnorm(100, mean = 0, sd = 1)      # 100 draws from N(0,1)
rbinom(50, size = 10, prob = 0.3)    # 50 draws from Binom(10,
    0.3)
rpois(100, lambda = 5)              # 100 draws from Poisson(5)
runif(100, min = 0, max = 1)      # 100 draws from Uniform(0,1)
rexp(100, rate = 2)                 # 100 draws from Exp(2)
```

Pattern: r + distribution name (r for “random”)

The Four Functions: d, p, q, r

Every distribution has four functions:

```
# For Normal(0,1):
dnorm(0)          # d = density (PDF value at x=0)
pnorm(1.96)       # p = probability (CDF: P(X <= 1.96))
qnorm(0.975)      # q = quantile (inverse CDF)
rnorm(100)        # r = random samples
```

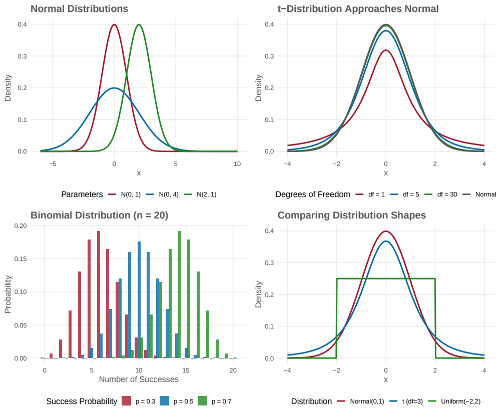
- d: “What’s the height of the density at this point?”
- p: “What’s the probability of being less than this?”
- q: “What value gives this probability?”
- r: “Give me random draws”

Visualizing Distributions

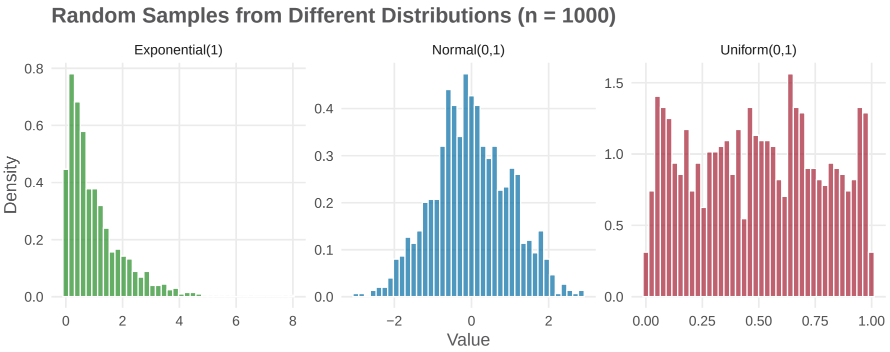
```r
library(ggplot2)

# Plot Normal PDF
x <- seq(-4, 4, length.out = 200)
ggplot(data.frame(x = x, y = dnorm(x)), aes(x, y)) +
  geom_line(color = "steelblue", linewidth = 1.2) +
  labs(title = "Standard Normal Distribution",
       x = "x", y = "Density")
```

Distribution Shapes



Sampling from Distributions



Random samples from different distributions (n = 1000 each).

Looking Ahead

Next week: Expected value and variance

- Defining $\mathbb{E}[X]$ and $\text{Var}[X]$ formally
- Properties: linearity, Chebyshev's inequality
- Covariance and correlation

Reading:

- Aronow & Miller, §2.1 (pp. 45–66)
- Blackwell, Chapter 2.4–2.5

Problem Set 1: Due February 14