

DISCRETE DISTRIBUTIONS

1. Bernoulli distribution

$X = 0$ or 1 (pass / fail)
 (ok / defective)
 (no error / error)
 (not infected / infected)
 (transmitted / lost)
Generic name: (successes / failures)

Called: *Bernoulli trial* (binary outcome)

X has Bernoulli pmf

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

So, there is a whole *family of Bernoulli distributions*,

$$\text{Bern}(p)$$

2. Binomial distribution

X = number of *successes* in n independent Bernoulli trials

Possible values: $X \in \{0, 1, \dots, n\}$

Binomial pmf: $p(x) = \binom{n}{x} p^x (1-p)^{n-x}$,

$x = 0, 1, \dots, n$

where

p^x = probability of x successes,

$(1-p)^{n-x}$ = probability of $(n-x)$ failures,

$\binom{n}{x} = \frac{n!}{x!(n-x)!}$ = number of outcomes

with exactly x successes and $(n-x)$ failures

There is a *family of Binomial distributions*,
 $Bin(n,p)$

n, p = parameters

n = number of trials

p = probability of success

Applications:

number of defective items in a sample

number of successful jobs

number of passing students

number of days without an accident

number of correct answers

Relation between **Bern(p)** and **Bin(n,p)**

If X_1, \dots, X_n are independent Bern(p) variables, then

$$X = \sum_{k=1}^n X_k \text{ is } Bin(n, p)$$

For $n = 1$, Bin(1,p) is Bern(p).

3. Geometric distribution

X_1, X_2, \dots = independent Bernoulli trials

X = the first successful trial
= number of trials needed
to see the first success

X has **Geometric** distribution with pmf

$$\begin{aligned} p(x) &= P\{X = x\} \\ &= P\{(x-1) \text{ failures, then 1 success}\} \end{aligned}$$

$$p(x) = (1-p)^{x-1}p$$

$$x = 1, 2, 3, \dots$$

Example: St. Petersburg paradox

4. Poisson distribution

X = number of “rare” events

(two events are unlikely to occur during a short period)

Examples:

Number of arrived jobs

Number of telephone calls

Number of claims

Number of errors

Number of accidents

Number of customers

Poisson distribution

Poisson pmf:

$$p(x) = e^{-\lambda} \frac{\lambda^x}{x!}$$

$$x = 0, 1, 2, \dots$$

λ = intensity parameter, average number of events per unit of time

Poisson approximation of the Binomial distribution

If p is small (< 0.05)
and n is large (≥ 30)
so that $np = \lambda$, then

$$Bin(n, p) \approx Poisson(\lambda)$$

If $p > 0.95$ then $(1 - p) < 0.05$, and

$$Bin(n, 1 - p) \approx Poisson(n(1 - p))$$

Mathematically,

$$\left\{ \begin{array}{l} \lim \\ p \rightarrow 0 \\ n \rightarrow \infty \\ np \rightarrow \lambda \end{array} \right\} \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} = e^{-\lambda} \frac{\lambda^x}{x!}$$

Examples

$n = 2$ mln. people in DFW area

$$p = P \left\{ \begin{array}{c} \text{a person calls NY} \\ \text{today} \end{array} \right\} = 7.5 \cdot 10^{-5}$$

$$\lambda = np = 150$$

Then $X =$ number of calls from DFW to NY
is

$$Bin(2 \cdot 10^6, 7.5 \cdot 10^{-5}) \approx Poisson(150)$$

Poisson distribution

X = number of accidents during rush hours

$n = 10800$ seconds during 3 evening rush hours

$p = 0.00015$ = probability of an accident during any given second

Then

X is *Poisson*(λ),

where $\lambda = np = 1.62$, the average number of accidents

Summary of Discrete Distributions

Distribution family	Description	Possible values	Parameters
Bernoulli	0 or 1, success or failure	0, 1	$p =$ prob. of success
Binomial	Number of successes in n Bernoulli trials	$0, \dots, n$	$n =$ number of trials $p =$ prob. of success
Geometric	Number of trials until success	$1, 2, \dots$	$p =$ prob. of success
Poisson	Number of rare events	$0, 1, \dots$	$\lambda =$ frequency of events

Summary of Discrete Distributions

Distribution family	P.M.F. $P(x)$	$\mathbf{E}(X)$	$\mathbf{Var}(X)$
Bernoulli	$P(0) = 1 - p$ $P(1) = p$	p	$p(1 - p)$
Binomial	$\binom{n}{x} p^x (1 - p)^{n-x}$ for $x = 0, 1, \dots, n$	np	$np(1 - p)$
Geometric	$(1 - p)^{x-1} p$ for $x = 1, 2, \dots$	$\frac{1}{p}$	$\frac{1 - p}{p^2}$
Poisson	$e^{-\lambda} \frac{\lambda^x}{x!}$ for $x = 0, 1, \dots$	λ	λ