

# 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*,  $Bin(n,p)$  (Table C.1)

$n, p$  = 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. Negative Binomial distribution

$X_1, X_2, \dots$  = independent Bernoulli trials

$X$  = the  $r$ -th successful trial  
= number of trials needed  
to see the  $r$ -th success

$X$  has **Negative Binomial** distribution with pmf

$$\begin{aligned} p(x) &= P\{X = x\} \\ &= P \left\{ \begin{array}{l} \text{Among } x \text{ trials: } r \text{ successes} \\ \text{and } (x - r) \text{ failures, and} \\ \text{the last trial is a success} \end{array} \right\} \end{aligned}$$

$$p(x) = \binom{x-1}{r-1} (1-p)^{x-r} p^r$$

$$x = r, r + 1, r + 2, \dots$$

## Relations

---

### Binomial variable

$n$  trials. How many successes?

### Negative Binomial variable

$r$  successes. How many trials?

---

For  $r = 1$ ,  $NegBin(1,p) = Geom(p)$

If  $X_1, \dots, X_r$  are  $Geom(p)$ , then

$$X = \sum_{i=1}^r X_i \text{ is } NegBin(r,p)$$

## 5. 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:

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

$x = 0, 1, 2, \dots$  (Table C.2)

$\lambda$  = 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 ( $\geq 30$ )  
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*( $\lambda$ ),

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