~~How to find a frequency function?~~

~~$p(k) = P(\{X=k\}) = \frac{\#\text{ of outcomes in } \{X=k\}}{\#\text{ of outcomes in } \Omega}$~~

~~Example: Hypergeometric distribution.~~

Instead we use Bernoulli trials

Discrete Distributions

A simplest example of random experiment is a coin-tossing, formally called "Bernoulli trial." It happens to be the case that many useful distributions are built upon this simplest form of experiment, whose relations are summarized later in a diagram.

Bernoulli trials.

A Bernoulli random variable X takes value only on 0 and 1. It is determined by the **parameter** p (which represents the probability that $X = 1$), and the frequency function is given by

$$P(X=1) = p(1) = p = P(X=1)$$

$$P(X=0) = p(0) = 1 - p = P(X=0) = 1 - P(X=1) = 1 - p$$

a subset of Ω

If A is the event that an experiment results in a "success," then the **indicator random variable**, denoted by I_A, takes the value 1 if A occurs and the value 0 otherwise.

$X : \Omega \ni \omega \to X(\omega) \in \mathbb{R}$

$$X(\omega) = I_A(\omega) = \begin{cases} 1 & \text{if } \omega \in A; \\ 0 & \text{otherwise (i.e., } \omega \notin A). \end{cases}$$

Then I_A is a Bernoulli random variable with "success" probability $p = P(A)$. We will call such experiment a **Bernoulli trial**.

Binomial distribution.

$\underbrace{X_1, X_2, \ldots, X_n}$ are independent Bernoulli trials (Bernoulli random variables) with success probability p

$\boxed{X = \# \text{ of successes}}$ ← A new random variable

$p(k) = P(X=k) = ?$

If we have n independent Bernoulli trials, each with a success probability p, then the probability that there will be exactly k successes is given by

$$\boxed{p(k) = \binom{n}{k} p^k (1-p)^{n-k}, \quad k = 0, 1, \ldots, n.}$$

← Binomial frequency function

The above frequency function $p(k)$ is called a **binomial distribution** with parameter (n, p).

Binomial Theorem

$$\boxed{(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}}$$

$$\sum_{k=0}^{n} p(k) = \sum_{k=0}^{n} \underbrace{\binom{n}{k} \underset{a^k}{\underset{\shortparallel}{p^k}} \underset{b^{n-k}}{\underset{\shortparallel}{(1-p)^{n-k}}}} = [\underset{a}{\underset{\shortparallel}{p}} + \underset{b}{\underset{\shortparallel}{(1-p)}}]^n = 1^n = 1$$

$\boxed{n=5}$, $\boxed{k=2}$

independent Bernoulli trials

$$\underbrace{X_1 \quad X_2 \quad X_3 \quad X_4 \quad X_5}_{\text{independent Bernoulli trials}} \qquad X = \#\text{ of successes}$$

$$\binom{5}{2}\left\{\begin{array}{ccccc|c}
\underset{1st}{\boxed{1}} & \underset{2nd}{\boxed{1}} & 0 & 0 & 0 & 2 \\
\underset{1}{\boxed{1}} & 0 & \underset{3}{\boxed{1}} & 0 & 0 & 2 \\
 & & \vdots & & & \\
0 & 0 & 0 & \underset{4}{\boxed{1}} & \underset{5th}{\boxed{1}} & 2
\end{array}\right.$$

$$P(X_1=1, X_2=1, X_3=X_4=X_5=0)$$
$$= \underline{P(X_1=1)}\,\underline{P(X_2=1)}\,\underline{P(X_3=0)}\,\underline{P(X_4=0)}\,\underline{P(X_5=0)}$$
$$= \underline{(p)}(p)\underline{(1-p)}(1-p)(1-p) = \underline{p^2(1-p)^3}$$

How many different ways to get exactly two successes?

‖

How many different way to pick two numbers from $\{1,2,3,4,5\}$?

$\rightarrow \boxed{\binom{5}{2}}$

$$P(X=2) = \binom{\overset{\nearrow n}{5}}{\underset{\searrow k}{2}} p^{\overset{k}{2}}(1-p)^{\overset{n-k}{3}} \quad \longleftarrow \text{ This is the case when } \boxed{n=5},\ \boxed{k=2}$$

↓ Generalize

$$\boxed{P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}}$$

Example

$n=5$

Five fair coins are flipped independently. Find the frequency function of the number of heads obtained.

$p=\frac{1}{2}$

$$p(k)=\binom{5}{k}\left(\frac{1}{2}\right)^{k}\left(1-\frac{1}{2}\right)^{5-k}=\binom{5}{k}\left(\frac{1}{2}\right)^{5}$$

Example

Five fair coins are flipped independently. Find the frequency function of the number of heads obtained.

The number X of heads represents a binomial random variable with parameter $n = 5$ and $p = \frac{1}{2}$. Thus, we obtain

$$p(k) = \binom{5}{k}\left(\frac{1}{2}\right)^5$$

Example

A company has known that their screws is defective with probability $p=0.01$. They sell the screws in packages of $\underset{n}{10}$, and are planning a money-back guarantee

1. at most one of the 10 screws is defective, and they replace it if a customer find more than one defective screws, or
2. they replace it even if there is only one defective.

For each of the money-back guarantee plans above what proportion of packages sold must be replaced?

$$p(k) = \binom{10}{k}(0.01)^k \underbrace{(1-0.01)}_{0.99}{}^{10-k}$$

① Let X be # of defective screws.

$$P(X \geq 2) = 1 - P(X \leq 1) = 1 - [p(0) + p(1)]$$

② $P(X \geq 1) = 1 - P(X = 0) = 1 - p(0)$

Example

A company has known that their screws is defective with probability 0.01. They sell the screws in packages of 10, and are planning a money-back guarantee

1. at most one of the 10 screws is defective, and they replace it if a customer find more than one defective screws, or
2. they replace it even if there is only one defective.

For each of the money-back guarantee plans above what proportion of packages sold must be replaced?

The number X of defective screws in a package represents a binomial random variable with parameter $p = 0.01$ and $n = 10$.

1. $P(X \geq 2) = 1 - p(0) - p(1) =$
   $1 - \binom{10}{0}(0.01)^0(0.99)^{10} - \binom{10}{1}(0.01)^1(0.99)^9 \approx \boxed{0.004}$
2. $P(X \geq 1) = 1 - p(0) = 1 - \binom{10}{0}(0.01)^0(0.99)^{10} \approx \boxed{0.096}$

$p(0) + \cdots + p(k) = P(X \leq k) = F(k)$ is cdf.

Properties of binomial random distributions.

By applying the binomial theorem, we can immediately see the following results:

$$\sum_{k=0}^{n} p(k) = \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \overset{\text{Binomial Theorem}}{=} [p + (1-p)]^n = 1;$$

$$M(t) = \sum_{k=0}^{n} \underbrace{(e^{kt})}_{g(k)} p(k) = \sum_{k=0}^{n} \binom{n}{k} (\underbrace{pe^t}_{a})^k (\underbrace{1-p}_{b})^{n-k} = \underline{[pe^t + (1-p)]^n}.$$

$E[e^{tX}]$ $\uparrow M(t)$

$\parallel$

$$\boxed{\begin{array}{c} E[g(X)] \\ \parallel \\ \sum_{k=0}^{n} g(k)\,p(k) \end{array}}$$

$\uparrow p(k)$: $\binom{n}{k} p^k (1-p)^{n-k}$

$\leftarrow$ Why do we care? $\leftarrow$

$M'(t)|_{t=0} = E[Xe^{tX}]|_{t=0} = E[X]$

$M''(t)|_{t=0} = E[X^2] \longrightarrow \mathrm{Var}(X) = E[X^2] - (E[X])^2$

Moments of binomial random distributions.

The mgf $M(t)$ allows us to calculate the moments $E[X]$ and $E[X^2]$.

$$M'(t) = \frac{d}{dt}\left([pe^t + (1-p)]^n\right) = npe^t[pe^t + (1-p)]^{n-1}$$

$$\begin{aligned} M''(t) &= \frac{d}{dt}\left(npe^t[pe^t + (1-p)]^{n-1}\right) \\ &= npe^t[pe^t + (1-p)]^{n-1} + n(n-1)(pe^t)^2[pe^t + (1-p)]^{n-2} \end{aligned}$$

Thus, we obtain $E[X] = \underline{M'(0)} = \boxed{\underline{np}}$ and $E[X^2] = M''(0) = np + n(n-1)p^2 = \underline{np + (np)^2 - np^2}$. This allows us to calculate $\mathrm{Var}(X) = E[X^2] - (E[X])^2 = \underline{np - np^2} = \underline{np(1-p)}$.

$$E[X] = \sum_{k=0}^{n} k\,p(k) = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k} = \boxed{np}$$

$$E[X^2] = \sum_{k=0}^{n} k^2 p(k) = np + (np)^2 - np^2 \longrightarrow \mathrm{Var}(X) = \boxed{np(1-p)}$$

Relation between Bernoulli trials and binomial random variable.

Recall: We begin with Bernoulli trials $x_1, \cdots, x_n \Rightarrow X = \#$ of successes

$$Y = x_1 + \cdots + x_n = \# \text{ of } 1\text{'s} = \# \text{ of successes} = X \sim \text{Binomial}$$

$$x_i = \begin{cases} \textcircled{1} & \text{if success} \rightarrow p(1) = p \\ \textcircled{0} & \text{if failure} \rightarrow p(0) = 1-p \end{cases}$$

A binomial random variable can be expressed in terms of n Bernoulli random variables. If $X_1, X_2, \ldots, X_n$ are independent Bernoulli random variables with success probability p, then the sum of those random variables

Claim

$$\boxed{Y = \sum_{i=1}^{n} X_i}$$

is **distributed as** the binomial distribution with parameter (n, p).

Their relationship in mgf.

Let X be a Bernoulli random variable with success probability p. Then the mgf $M_X(t)$ is given by

$$M_X(t) = e^{(0)t} \times \underbrace{(1-p)}_{p(0)} + e^{(1)t} \times \underbrace{p}_{p(1)} = \underline{(1-p) + pe^t}.$$

$$= \sum_{k=0}^{1} e^{kt} p(k)$$

Then we can derive the mgf of a binomial random variable $Y = \sum_{i=1}^{n} X_i$ from the mgf of Bernoulli random variable by applying the mgf property for independent random variables repeatedly.

$$M_Y(t) = M_{X_1}(t) \times M_{X_2}(t) \times \cdots \times M_{X_n}(t) = \underbrace{[(1-p) + pe^t]^{\textcircled{n}}}_{\text{mgf for binomial distr.}}.$$

$$= E[e^{tY}]$$

$$= E[e^{t\underbrace{(X_1+\cdots+X_n)}_{\sum_{i=1}^{n} X_i}}] = E[e^{tX_1} \cdot e^{tX_2} \cdot \cdots \cdot e^{tX_n}] \underset{X_1, \ldots, X_n \text{ are independent}}{=} \overset{M(t)}{E[e^{tX_1}]} \cdot \overset{M(t)}{E[e^{tX_2}]} \cdot \cdots \cdot \overset{M(t)}{E[e^{tX_n}]}$$

Sum of two independent random variables.

Theorem

If X and Y are independent binomial random variables with respective parameters (n, p) and (m, p), then the sum $Z = X + Y$ is distributed as the binomial distribution with parameter $(n + m, p)$.

1st attempt of proof.

X = # of successes in the first n trials

+) Y = " the next m trials

Z = " in $(n+m)$ trials ~ Binomial with $(n+m, p)$

Sum of two independent random variables.

Theorem

If X and Y are independent binomial random variables with respective parameters (n, p) and (m, p), then the sum $Z = X + Y$ is distributed as the binomial distribution with parameter $(n + m, p)$.

By the mgf property for independent random variables we obtain

$$M_Z(t) \overset{Z=X+Y}{=} \underline{M_X(t)} \times \underline{M_Y(t)} = \underline{[(1-p)+pe^t]}^{\textcircled{n}} \times \underline{[(1-p)+pe^t]}^{\textcircled{m}}$$

$$= [(1-p)+pe^t]^{n+m} \rightarrow \text{Binomial MGF.}$$

This uniquely determines the distribution for Z, which is binomial with parameter $(n + m, p)$.

$$M_Z(t) = E[e^{tZ}] = E[e^{tX} \cdot e^{tY}]$$

$$= E[e^{tX}] \cdot E[e^{tY}]$$

Alternative approach to the proof.

$$\underbrace{Z_1, \dots, Z_n}, \underbrace{Z_{n+1}, \dots, Z_{n+m}}$$

Observe that we can express

$$X = \sum_{i=1}^{n} Z_i \text{ and } Y = \sum_{i=n+1}^{n+m} Z_i$$

in terms of independent Bernoulli random variables Z_i's with success probability p. Thus, the resulting sum

$$Z = X + Y = \sum_{i=1}^{n+m} Z_i$$

must be a binomial random variable with parameter $(n + m, p)$.

Expectations of Bernoulli and binomial random variables.

$$\left\{ \begin{array}{l} p_X(1) = \boxed{p} \\ p_X(0) = \boxed{1-p} \end{array} \right.$$

Let X be a Bernoulli random variable with success probability p. Then the expectation of X becomes

$$\underset{= \sum_{k=0} k\,p(k)}{E}[X] = 0 \times \underbrace{(1-p)}_{p(0)} + 1 \times \underbrace{p}_{p(1)=p} = \boxed{p.}$$

Now let Y be a binomial random variable with parameter (n, p). Recall that $\boxed{Y}$ can be expressed as the sum $\boxed{\sum_{i=1}^{n} X_i}$ of independent Bernoulli random variables $X_1, \ldots, X_n$ with success probability p. Thus, by using property (c) of expectation we obtain

$$\underset{\text{Binomial r.v.}}{\underset{\uparrow}{E[Y]}} = E\left[\sum_{i=1}^{n} \underset{\text{Bernoulli r.v.'s}}{\underset{\uparrow}{X_i}}\right] = \sum_{i=1}^{n} \underbrace{E[X_i]}_{p} = \boxed{np.}$$

Variances of Bernoulli and binomial random variables.

$$\begin{cases} p(0) = 1-p \\ p(1) = p \end{cases}$$

Let X be a Bernoulli random variable with success probability p. Then the expectation $E[X]$ is p, and the variance of X is

$$E[(X-E[X])^2] = \sum_{k=0}^{1} (k-\overset{\downarrow}{p})^2 p(k)$$

$$\mathrm{Var}(X) = (\underbrace{0}_{k} - p)^2 \times (\overbrace{1-p}^{p(0)}) + (\underbrace{1}_{k} - p)^2 \times \overset{p(1)}{p} = p(1-p).$$

$$p^2(1-p) + (1-p)^2 p = (1-p)p[p+(1-p)]$$

Since a binomial random variable Y is the sum $\sum_{i=1}^{n} X_i$ of independent Bernoulli random variables, we obtain

$$\mathrm{Var}(\overset{\text{Binomial r.v}}{Y}) = \mathrm{Var}\left(\sum_{i=1}^{n} \overset{\text{Bernoulli r.v's}}{X_i}\right) = \sum_{i=1}^{n} \underbrace{\mathrm{Var}(X_i)}_{p(1-p)} = np(1-p).$$

$$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\underbrace{\mathrm{Cov}(X,Y)}_{0 \text{ if } X \text{ and } Y \text{ are independent.}}$$

$$\Downarrow$$

$$\mathrm{Var}(X_1+X_2+X_3) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2) + \mathrm{Var}(X_3)$$

Hypergeometric distribution.

← Recall from note#1

Consider the collection of N subjects, of which m belongs to one particular class (say, "tagged" subjects), and $(N-m)$ to another class (say, "non-tagged"). Now a sample of size r is chosen randomly from the collection of N subjects. Then the number X of "tagged" subjects selected has the frequency function

Pick r samples

X = # of tagged subjects out of the sample of size r.

$$p(k) = \frac{\binom{m}{k}\binom{N-m}{r-k}}{\binom{N}{r}}$$

E(x) and Var(x)?

where $0 \leq k \leq r$ must also satisfy $k \geq r-(N-m)$ and $k \leq m$. The above frequency function $p(k)$ is called a **hypergeometric** distribution with parameter (N, m, r).

$$P(X=k) = \frac{(\#\text{ of outcomes in } \{X=k\})}{(\quad " \qquad \text{in } \Omega)}$$

$$\frac{\binom{m}{k}\times\binom{N-m}{r-k}}{\binom{N}{r}}$$

Example

A lot, consisting of 50 items, is inspected. Suppose that the lot contains 5 defective items. In the following two inspection procedures, What is the probability that no defective items are found in the inspection?

1. Four items are randomly selected and tested.
2. Ten items are randomly selected and tested.

X = # of defective items ~ Hypergeometric distribution $N=50$, $m=5$, $r=\begin{cases} 4 & \text{in (1)} \\ 10 & \text{in (2)} \end{cases}$

(1) $P(X=0) = \dfrac{\binom{m}{0}\binom{N-m}{r-0}}{\binom{N}{r}}$

Example

A lot, consisting of 50 items, is inspected. Suppose that the lot contains 5 defective items. In the following two inspection procedures, What is the probability that no defective items are found in the inspection?

1. Four items are randomly selected and tested.
2. Ten items are randomly selected and tested.

The number X of defective items in the inspection has a hypergeometric distribution with $N = 50$, $m = 5$.

1. Here we choose $r = 4$, and calculate $P(X = 0) = \frac{\binom{5}{0}\binom{45}{4}}{\binom{50}{4}} \approx 0.65$
2. Here we choose $r = 10$, and calculate $P(X = 0) = \frac{\binom{5}{0}\binom{45}{10}}{\binom{50}{10}} \approx 0.31$

Relation between Bernoulli trials and Hypergeometric distribution.

What is $P(A_i)$

Let A_i be the event that a “tagged” subject is found at the *i*-th selection.

$$X_i = \begin{cases} 1 & \text{if } A_i \text{ occurs;} \\ 0 & \text{otherwise.} \end{cases}$$

Pick r subjects : $x_1 = \begin{cases} 1 \\ 0 \end{cases}, \cdots, x_r = \begin{cases} 1 \\ 0 \end{cases}$

Why ?

Then X_i is a Bernoulli trial with “success” probability $p = P(A_i) = \frac{m}{N}$. Then the number Y of “tagged” subjects in a sample of size r can be expressed in terms of r Bernoulli random variables.

$$Y = \sum_{i=1}^{r} X_i \longrightarrow E[Y] = \sum_{i=}^{r} \underbrace{E[X_i]}_{p} = rp = r\left(\frac{m}{N}\right)$$

Hypergeometric

Bernoulli r.v's

is **distributed as** a hypergeometric distribution with parameter (N, m, r).

Example

tagged subject

An urn contains 10 red balls and 20 blue balls, and balls are drawn one at a time without replacement. Let A_i be the event that a red ball is drawn at the i-th attempt.

1. Find $P(A_1)$ $= \frac{10}{30} = \boxed{\frac{1}{3}}$
2. Calculate $P(A_2)$. $= \frac{\#\text{ of outcomes in } A_2}{\#\text{ of outcomes in } \Omega}$ $\frac{24 \times 10}{30 \times 29} = \boxed{\frac{1}{3}}$
3. Find $P(A_i)$ in general.

$$P(A_i) = \frac{\#\text{ of outcomes in } A_i}{\#\quad '' \quad \text{in } \Omega} = \frac{\cancel{29 \times 28 \times \cdots \times (30-i+1)} \times 10}{30 \times \cancel{29 \times \cdots \times (30-i+1)}} = \frac{1}{3}$$

1st 2nd ith

Example

An urn contains 10 red balls and 20 blue balls, and balls are drawn one at a time without replacement. Let A_i be the event that a red ball is drawn at the i-th attempt.

1. Find $P(A_1)$
2. Calculate $P(A_2)$.
3. Find $P(A_i)$ in general.

1. $P(A_1) = \frac{10}{30} = \frac{1}{3}$.
2. If we draw 2 balls then we have 10 × 29 outcomes in which the second ball is red. Thus, $P(A_2) = \frac{10 \times 29}{30 \times 29} = \frac{1}{3}$.
3. In general we have $10 \times 29 \times 28 \times \cdots \times (31 - i)$ outcomes in which i-th ball is red, and obtain

$$P(A_i) = \frac{\overset{i\text{th place}}{\boxed{10}} \times \cancel{29 \times 28 \times \cdots \times (31 - i)}}{30 \times \cancel{29 \times 28 \times \cdots \times (31 - i)}} = \frac{1}{3}.$$

Expectation of hypergeometric random variable.

Let X_i be a Bernoulli trial of finding a "tagged" subject in the i-th selection. Then the expectation of X becomes

$$E[X_i] = \frac{m}{N}. = p \quad \text{for each } i$$

Now let Y be a hypergeometric random variable with parameter (N, m, r). Recall that $\boxed{Y}$ can be expressed as the sum $\boxed{\sum_{i=1}^{r} X_i}$ of the above Bernoulli trials. We can easily calculate

$$\sum_{k=0}^{r} k \frac{\binom{m}{k}\binom{N-m}{r-k}}{\binom{N}{r}} = E[Y] = E\left[\sum_{i=1}^{r} X_i\right] = \sum_{i=1}^{r} \underbrace{E[X_i]}_{\frac{m}{N}} = \boxed{\frac{mr}{N}} = rp$$

$p = \frac{m}{N}$

How about $\mathrm{Var}(Y)$? Is it true: $\mathrm{Var}(Y) = \sum_{i=1}^{r} \mathrm{Var}(X_i)$ ~~because of independence~~ ?

No !

Dependence of Bernoulli trials in Hypergeometric distribution.

Suppose that $i \neq j$. Then we can find that

$$X_iX_j = \begin{cases} 1 & \text{if } A_i \text{ and } A_j \text{ occurs;} \\ 0 & \text{otherwise.} \end{cases}$$

is again a Bernoulli trial with "success" probability $p = P(A_i \cap A_j) = \frac{m(m-1)}{N(N-1)}$. Since $E[X_i] = E[X_j] = \frac{m}{N}$ $= p$ and $E[X_iX_j] = \frac{m(m-1)}{N(N-1)}$ $= p$, we can calculate

$$\mathrm{Cov}(X_i, X_j) = \underbrace{E[X_iX_j]}_{\frac{m(m-1)}{N(N-1)}} - \underbrace{E[X_i]}_{\frac{m}{N}}\underbrace{E[X_j]}_{\frac{m}{N}} = \frac{m(m-N)}{N^2(N-1)} < 0$$

Therefore, X_i and X_j are dependent, and negatively correlated.

$$\frac{m(m-1)}{N(N-1)} - \left(\frac{m}{N}\right)^2 = \frac{m}{N^2(N-1)}\left[\underbrace{N(m-1) - m(N-1)}_{m-N}\right]$$

Recall : 10 red ball and 20 blue ball.

$$P(A_1 \cap A_2) = \frac{10 \times 9}{30 \times 29}$$

$$= \frac{m(m-1)}{N(N-1)}$$

$$P(A_i \cap A_j) = \frac{m(m-1)}{N(N-1)}$$

$$= P(A_1 \cap A_2)$$

$$\mathrm{Cov}(X, Y)$$

$$= E[XY] - E[X]E[Y]$$

Variance of hypergeometric random variable.

The variance of Bernoulli trial with success probability $\overset{P}{\frac{m}{N}}$ is given by

$$\mathrm{Var}(X_i) = \left(\frac{m}{N}\right)\left(1-\frac{m}{N}\right) = \frac{m(N-m)}{N^2}$$

$p = \frac{m}{N}$ $= p(1-p)$

Together with $\mathrm{Cov}(X_i, X_j) = \frac{m(m-N)}{N^2(N-1)}$, we can calculate

How many pairs?

$\binom{r}{2} = \frac{r(r-1)}{2}$

All the different pairs of (i, j)

$1 \le i < j \le r$

$$\mathrm{Var}(Y) = \mathrm{Var}\left(\sum_{i=1}^{r} X_i\right) = \sum_{i=1}^{r} \mathrm{Var}(X_i) + 2\sum_{i<j} \mathrm{Cov}(X_i, X_j)$$

Hypergeometric

$$= r \times \frac{m(N-m)}{N^2} + r(r-1) \times \frac{m(m-N)}{N^2(N-1)}$$

$$= \frac{mr(N-m)(N-r)}{N^2(N-1)}$$

$\mathrm{Var}(X_1+X_2) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2) + 2\,\mathrm{Cov}(X_1, X_2)$

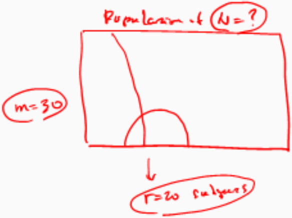
$\mathrm{Var}(X_1+X_2+X_3) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2) + \mathrm{Var}(X_3) + 2\,\mathrm{Cov}(X_1, X_2)$
$+ 2\,\mathrm{Cov}(X_1, X_3)$
$+ 2\,\mathrm{Cov}(X_2, X_3)$

It is generalized

Example

We capture 30 individuals, mark them with a yellow tag, and then release them back into the environment. Later we capture another sample of 20 individuals. Some of the individuals in this sample have been marked, known as "recaptures." Let X be the number of recaptures. Suppose that the population size N is unknown.

1. What is the frequency function for X?
2. How can we guess the population size N from the observed number X of recaptures?



$$2 = X \approx E[X] = \frac{rm}{N} = \frac{(20)(30)}{N}$$

$$\Downarrow$$

$$\boxed{N \approx \frac{(20)(30)}{2}}$$

X = # of tagged subjects ~ Hypergeometric

Example

We capture 30 individuals, mark them with a yellow tag, and then release them back into the environment. Later we capture another sample of 20 individuals. Some of the individuals in this sample have been marked, known as “recaptures.” Let X be the number of recaptures. Suppose that the population size N is unknown.

1. What is the frequency function for X?
2. How can we guess the population size N from the observed number X of recaptures?

1. It is hypergeometric with $m = 30$ and $r = 20$, which gives $p(k) = \frac{\binom{30}{k}\binom{N-30}{20-k}}{\binom{N}{r}}$ for $k = 0, 1, \ldots, 20$, assuming that $N \geq 50$.
2. Since the best guess for $\overset{\text{observed}}{\enclose{circle}{X}}$ is the mean value $E[X] = \frac{mr}{N} = \frac{600}{N}$, we can form the equation $\boxed{X \approx \frac{600}{N}}$ in approximation. By solving it, we can estimate $\enclose{circle}{N} \approx \enclose{circle}{\frac{600}{X}}$ if $X \geq 1$. If $X = 0$, there is no guess for N.

Geometric series formulas.

For $0 < a < 1$, we can establish the following formulas.

$$\sum_{k=1}^{n} a^{k-1} = \frac{1-a^{n}}{1-a}; \quad a^{n} \to 0 \quad \xrightarrow{n\to\infty} \quad \sum_{k=1}^{\infty} a^{k-1} = \frac{1}{1-a}; \tag{3.1}$$

Assume $0 < x < 1$

$$\sum_{k=1}^{\infty} k a^{k-1} = \frac{d}{dx}\left(\sum_{k=0}^{\infty} x^{k}\right)\Bigg|_{x=a} = \frac{d}{dx}\left(\frac{1}{1-x}\right)\Bigg|_{x=a} = \frac{1}{(1-a)^2}; \tag{3.2}$$

$$\sum_{k=1}^{\infty} k^{2} a^{k-1} = \frac{d}{dx}\left(\sum_{k=1}^{\infty} k x^{k}\right)\Bigg|_{x=a} = \frac{d}{dx}\left(\frac{x}{(1-x)^2}\right)\Bigg|_{x=a}$$

$$= \frac{1}{(1-a)^2} + \frac{2a}{(1-a)^3}$$

$$\frac{d}{dx} x^{k} = k x^{k-1}$$

$$\sum_{k=1}^{\infty} k x^{k} = x\left(\sum_{k=1}^{\infty} k x^{k-1}\right) = x\left(\frac{1}{(1-x)^2}\right)$$

Geometric random variables.

X_1 X_2 X_3 X_4 X_5 . . . . .

0 ① 0 0 1 → $X = 2$

0 0 0 ① 0 → $X = 4$

X = # of Bernoulli trials until 1st success

$P(X=2) = (1-p)p$, $P(X=4) = (1-p)(1-p)(1-p)p = (1-p)^3 p \xrightarrow{\text{In general}} P(X=k) = (1-p)^{k-1} p$

In sampling independent Bernoulli trials, each with probability of success p, the frequency function for the number of trials at the first success is

X

$$p(k) = (1-p)^{k-1}p \qquad k = 1, 2, \ldots .$$

This is called a *geometric distribution*. A random variable X is called a *geometric random variable* if it has a frequency function

$$P(X = k) = (1-p)^{k-1}p$$

for $k = 1, 2, \ldots$.

$$\sum_{k=1}^{\infty} p(k) = \sum_{k=1}^{\infty} (1-p)^{k-1} p = p \sum_{k=1}^{\infty} \underbrace{(1-p)^{k-1}}_{a^{k-1}} = p \frac{1}{(1-(1-p))} = p \cdot \frac{1}{p} = 1$$

Example

Consider a roulette wheel consisting of 38 numbers—1 through 36, 0, and double 0. If Mr.Smith always bets that the outcome will be one of the numbers 1 through 12, what is the probability that (a) his first win will occur on his fourth bet; (b) he will lose his first 5 bets?

X = # of bets until he wins at the first time

$$X_i = \begin{cases} 1 & \text{if he wins at } i\text{th bet} \\ 0 & \text{if he lose } \text{"} \end{cases}$$

success probability $= \frac{12}{38} = \frac{6}{19} = p$

failure probability $= \frac{13}{19} = 1 - \frac{6}{19} = 1 - p$

(a) $$P(X=4) = \underbrace{\left(\frac{13}{19}\right)\left(\frac{13}{19}\right)\left(\frac{13}{19}\right)}_{3 \text{ bets lost}}\left(\frac{6}{19}\right) = \left(\frac{13}{19}\right)^3\left(\frac{6}{19}\right) = (1-p)^3 p$$

$$\vdots$$

$$P(X=k) = \underbrace{\left(\frac{13}{19}\right)\cdots\left(\frac{13}{19}\right)}_{(k-1) \text{ bets lost}}\left(\frac{6}{19}\right) = \left(\frac{13}{19}\right)^{k-1}\left(\frac{6}{19}\right) = (1-p)^{k-1} p$$

(b) $$P(\text{he lose 5 bets in a row}) = \left(\frac{13}{19}\right)\left(\frac{13}{19}\right)\left(\frac{13}{19}\right)\left(\frac{13}{19}\right)\left(\frac{13}{19}\right) = \left(\frac{13}{19}\right)^5$$

Example

Consider a roulette wheel consisting of 38 numbers—1 through 36, 0, and double 0. If Mr.Smith always bets that the outcome will be one of the numbers 1 through 12, what is the probability that (a) his first win will occur on his fourth bet; (b) he will lose his first 5 bets?

The number X of bets until the first win is a geometric random variable with success probability $p = \left(\frac{12}{38}\right) = \left(\frac{6}{19}\right)$.

$$\boxed{p(k) = (1-p)^{k-1}p}$$

$$\frac{13}{19}$$

(a) $P(X = 4) = \left(\frac{13}{19}\right)^{3}\left(\frac{6}{19}\right) \approx 0.1$

(b) $P(X \geq 6) = \sum_{k=6}^{\infty}(1-p)^{k-1}p$ $= \sum_{k=6}^{\infty} p(k) = (1-p)^5 p + (1-p)^6 p + \cdots$

$= (1-p)^5 p\,[1 + (1-p) + \cdots]$

$$\sum_{j=1}^{\infty} a^{j-1} = \frac{1}{1-a}$$

$$= (1-p)^5 p \underbrace{\sum_{j=1}^{\infty}(1-p)^{j-1}}_{\frac{1}{[1-(1-p)]} = \frac{1}{p}} = (1-p)^5 = \left(\frac{13}{19}\right)^5 \approx 0.15$$

$\{X \geq 6\} = \{\underbrace{\text{He loses the first five bets}}_{\left(\frac{13}{19}\right)^5}\}$; thus, $P(X \geq 6) = \left(\frac{13}{19}\right)^5$.

Properties of geometric distribution

By using the geometric series formula (3.1) we can immediately see the following:

$$\sum_{k=1}^{\infty} a^{k-1} = \frac{1}{1-a} \quad \text{if } \boxed{a<1}$$

$$\sum_{k=1}^{\infty} \overset{(1-p)^{k-1}p}{\overset{\parallel}{p(k)}} = p\sum_{k=1}^{\infty}(1-p)^{k-1} \overset{\downarrow}{=} \frac{p}{1-(1-p)} = 1;$$

$$M(t) = \sum_{k=1}^{\infty} e^{kt}p(k) = \boxed{pe^t}\sum_{k=1}^{n}[\underbrace{(1-p)e^t}_{\textcircled{a}}]^{k-1} = \frac{\boxed{pe^t}}{1-(1-p)e^t}.$$

$$= E[e^{tX}] \uparrow \qquad e^{kt} = e^t \cdot e^{(k-1)t}$$

$$\boxed{E[g(X)] = \sum_{k=1}^{\infty} g(k)\,p(k)}$$

Assume t is close to zero so that $a = \underline{(1-p)e^t} < 1$

Expectation and variance of geometric distribution.

Let X be a geometric random variable with success probability p. To calculate the expectation $E[X]$ and the variance $\mathrm{Var}(X)$, we apply the geometric series formula (3.2) to obtain

$$\sum_{k=1}^{\infty} k a^{k-1} = \frac{1}{(1-a)^2}$$

$$\sum_{k=1}^{\infty} k^2 a^{k-1} = \frac{1}{(1-a)^2} + \frac{2a}{(1-a)^3}$$

$$E[X] = \sum_{k=1}^{\infty} k\overset{p(1-p)^{k-1}}{\overset{\|}{p(k)}} = p \sum_{k=1}^{\infty} k(\underbrace{1-p}_{a})^{k-1} = \boxed{\frac{1}{p}};$$

$$\mathrm{Var}(X) = E[X^2] - (E[X])^2 = \sum_{k=1}^{\infty} k^2 p(k) - \left(\frac{1}{p}\right)^2$$

$$= p\left(\frac{1}{p^2} + \frac{2(1-p)}{p^3}\right) - \frac{1}{p^2} = \boxed{\frac{1-p}{p^2}}$$

Moments of geometric distribution.

$\underbrace{M'(0)}_{\parallel} = E[X], \quad M''(0) = E[X^2]$

$\frac{d}{dt}M(t)\Big|_{t=0}$

We can calculate $M'(t)$ and $M''(t)$ for $M(t) = \frac{pe^t}{1-(1-p)e^t}$. if t is close to zero

$$\frac{pe^t[1-(1-p)e^t] + pe^t(1-p)e^t}{[1-(1-p)e^t]^2}$$

$$M'(t) = \frac{pe^t}{(1-(1-p)e^t)^2}$$

$$M''(t) = \frac{d}{dt}\left[\frac{pe^t}{(1-(1-p)e^t)^2}\right] = \underline{\frac{pe^t + p(1-p)e^{2t}}{(1-(1-p)e^t)^3}}$$

Then we obtain $E[X] \underset{\uparrow_{t=0}}{=} M'(0) = \boxed{\frac{1}{p}}$ and $E[X^2] \underset{\uparrow_{\boxed{t=0}}}{=} M''(0) = \boxed{\frac{2-p}{p^2}}$. Thus, we have

$$\mathrm{Var}(X) = \underbrace{E[X^2]}_{\frac{2-p}{p^2}} - (E[X])^2 = \frac{1-p}{p^2}$$

Negative binomial distribution.

$r=3$

x = # of trials until r successes observed

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	
$\binom{4}{2}$	① 1st	0	① 3rd	0	①			$x=5$
	0	① 2nd	0	① 4th	①			$x=5$
	1	0	1	0	0	①		6

When r=3, $P(X=5) = (1-p)^2 \; p^3 \binom{4}{2} \xrightarrow{\text{In general}} P(X=k) = \binom{k-1}{\textcircled{2}}(1-p)^{k-\textcircled{3}} p^{\textcircled{3}}$

For general r

$P(X=k) = \binom{k-1}{r-1}(1-p)^{k-r}p^r$

Given independent Bernoulli trials with probability of success p, the frequency function of the number of trials until the r-th success is

$$p(k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}, \quad k = r, r+1, \ldots .$$

This is called a *negative binomial distribution* with parameter (r, p).

Example

$p = \frac{1}{2}$ $\quad$ $r = 4$

A fair coin is flipped repeatedly until heads are observed four times. Find the frequency function of the number of attempts to flip a coin.

X

$$P(X=k) = \binom{k-1}{3}\left(\frac{1}{2}\right)^{k-4}\left(\frac{1}{2}\right)^{4} = \binom{k-1}{3}\left(\frac{1}{2}\right)^{k} \longleftarrow p(k) = \binom{k-1}{r-1}p^{r}(1-p)^{k-r}$$

pick 3 places
where heads observed

$$\underbrace{① \; ② \; \cdots\cdots \; (k-1)}_{3 \text{ heads}} \quad \overset{\boxed{k}}{\underset{\text{4th heads}}{\uparrow}}$$

$[(k-1)-3] = (k-4)$ tails

Example

A fair coin is flipped repeatedly until heads are observed four times. Find the frequency function of the number of attempts to flip a coin.

The number X of attempts represents a negative binomial random variable with parameter $r = 4$ and $p = \frac{1}{2}$. Thus, we obtain

$$p(k) = \binom{k-1}{3} \left(\frac{1}{2}\right)^k, \quad k = 4, 5, 6, \ldots$$

Relation between geometric and negative binomial distribution.

$$\overbrace{Z_1\ Z_2\ \vdots\ Z_3\ Z_4\ Z_5\ \vdots\ Z_6\ Z_7\ \vdots\ Z_8\ Z_9}^{\text{Bernoulli trials}}$$

$$\underbrace{0\ \textcircled{1}}_{\boxed{X_1=2}}\ \vdots\ \underbrace{0\ 0\ \textcircled{1}}_{\boxed{X_2=3}}\ \vdots\ \underbrace{0\ \textcircled{1}}_{\boxed{X_3=2}}\ \vdots\ \underbrace{0\ \textcircled{1}}_{\boxed{X_4=2}}$$

Geometric r.v.s being independent

$\boxed{r=4}$

Y = # of trials until 4th success

$\textcircled{9}$

$$Y = X_1 + X_2 + X_3 + X_4 = \sum_{i=1}^{4} X_i$$

The geometric distribution is a special case of negative binomial distribution when $r = 1$. Moreover, if $X_1, \ldots, X_r$ are independent and identically distributed (iid) geometric random variables with parameter p, then the sum

\# of trials until r successes observed.

$$\boxed{Y} = \sum_{i=1}^{\boxed{r}} \boxed{X_i} \tag{3.3}$$

becomes a negative binomial random variable with parameter (r, p).

Expectation, variance and mgf of negative binomial distribution.

By using the sum of iid geometric random variables in (3.3), we can compute the expectation, the variance, and the mgf of negative binomial random variable Y.

$$\underline{E[Y]} = E\left[\sum_{i=1}^{r} X_i\right] = \sum_{i=1}^{r} \overbrace{E[X_i]}^{\left(\frac{1}{p}\right)} = \boxed{\frac{r}{p}};$$

$$\mathrm{Var}(Y) = \mathrm{Var}\left(\underline{\sum_{i=1}^{r} X_i}\right) = \sum_{i=1}^{r} \underbrace{\mathrm{Var}(X_i)}_{\left(\frac{1-p}{p^2}\right)} = \frac{r(1-p)}{p^2};$$

$$M_Y(t) = M_{X_1}(t) \times M_{X_2}(t) \times \cdots \times M_{X_r}(t) = \left[\frac{pe^t}{1-(1-p)e^t}\right]^{\boxed{r}}.$$

$X_1 + X_2 + \cdots + X_r$ $\nearrow M_Y$

$M_{X_1}(t) \Uparrow \left[\frac{pe^t}{1-(1-p)e^t}\right]$

$$M_{\underline{X_1+X_2}}(t) = E[e^{(X_1+X_2)t}] = E[\underbrace{e^{tX_1} \times e^{tX_2}}_{g_1(X_1) \times g_2(X_2)}] = E[e^{tX_1}]\,E[e^{tX_2}] = \underline{M_{X_1}}(t)\,\underline{M_{X_2}}(t)$$

if X_1 and X_2 are independent

Example

What is the average number of times one must throw a die until the outcome "1" has occurred 4 times?

$r=4$

Each attempt is a Bernoulli trial with success probability $p=\frac{1}{6}$

Y = # of trials until 4 successes

$$E[Y] = \frac{r}{p} = \frac{4}{\left(\frac{1}{6}\right)} = 24$$

Example

What is the average number of times one must throw a die until the outcome “1” has occurred 4 times?

The number X of throws until the fourth outcome of “1” is recorded is a negative binomial random variable with $r = 4$ and $p = \frac{1}{6}$. Thus, we obtain $E[X] = \frac{r}{p} = 24$.

Euler's number and natural exponential.

$$\boxed{\lim_{n\to\infty}\left(1+\frac{a}{n}\right)^n = e^a}$$

The number

$$e = \lim_{n\to\infty}\left(1+\frac{\textcircled{1}}{n}\right)^n = 2.7182\cdots = e^1$$

is known as the base of the natural logarithm (or, called Euler's number). The exponential function $f(x) = e^x$ is associated with the Taylor series

$$\underset{\sim\sim\sim\sim\sim\sim\sim}{e^x = \sum_{k=0}^{\infty}\frac{x^k}{k!}}. \tag{3.4}$$

Poisson distribution.

The **Poisson distribution** with parameter λ ($\lambda > 0$) has the frequency function

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

By applying the Taylor series (3.4) we can immediately see the following:

$$\sum_{k=0}^{\infty} p(k) = e^{-\lambda} \overset{\text{"}e^{\lambda}}{\boxed{\sum_{k=0}^{\infty} \frac{\lambda^k}{k!}}} = e^{-\lambda}e^{\lambda} = \enclose{circle}{1.}$$

A Poisson random variable X represents the number of successes when there are a large (and usually unknown) number of independent trials.

Example

Suppose that the number of typographical errors per page has a Poisson distribution with parameter $\lambda = 0.5$.

1. Find the probability that there are no errors on a single page.
2. Find the probability that there is at least one error on a single page.

X = # of typos in a page

$$P(X=k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

① $P(X=0) = p(0) = e^{-\lambda} \frac{\lambda^0}{0!} = e^{-\lambda} = e^{-0.5}$

② $P(X \geq 1) = 1 - P(X=0) = 1 - e^{-0.5}$

Example

Suppose that the number of typographical errors per page has a Poisson distribution with parameter $\lambda = 0.5$.

1. Find the probability that there are no errors on a single page.
2. Find the probability that there is at least one error on a single page.

1. $P(X = 0) = p(0) = e^{-0.5} \approx \boxed{0.607}$
2. $P(X \geq 1) = 1 - p(0) = 1 - e^{-0.5} \approx \boxed{0.393}$

Mgf and moments of Poisson distribution.

$e^{-\lambda}\frac{\lambda^k}{k!}$ $\qquad$ $e^x = \sum_{k=0}^{\infty}\frac{x^k}{k!}$

By applying the Taylor series (3.4) we obtain

$$M(t) = \sum_{k=0}^{\infty} \underbrace{e^{kt}}_{(e^t)^k} p(k) = e^{-\lambda}\sum_{k=0}^{\infty}\frac{(e^t\lambda)^k}{k!} = e^{-\lambda}e^{\lambda e^t} = \boxed{e^{\lambda(e^t-1)}}.$$

$= E[e^{tX}]$

Then we can calculate $M'(t) = \underline{\overset{g}{\lambda e^t}\,\overset{f}{e^{\lambda(e^t-1)}}}$ and $M''(t) = \underline{(\lambda e^t)^2 e^{\lambda(e^t-1)} + \lambda e^t e^{\lambda(e^t-1)}}$. Thus, we obtain $E[X] = M'(0) = \boxed{\lambda}$ and $E[X^2] = M''(0) = \lambda^2 + \lambda$. Finally we get $\underline{\mathrm{Var}(X) = E[X^2] - (E[X])^2} = \boxed{\lambda}$.

⇓

$\boxed{\lambda = E[X]} = Var(X) = \lambda$

↑ rate parameter

Poisson approximation to binomial distribution.

X = # of successes out of n trials ~ Binomial

$P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$ $\quad E[X] = np = \lambda$

average successes in Binomial

Poisson

The Poisson distribution can be used as an approximation for a binomial distribution with parameter (n, p) when n is large and p is small enough so that np is a moderate number λ. Let $\boxed{\lambda = np}$. Then the binomial frequency function becomes

as $n \to \infty$ we see $\boxed{p = \frac{\lambda}{n}} \to 0$

$$\underbrace{P(X=k) = p(k)}_{k \text{ is fixed.}} = \frac{n!}{k!(n-k)!}\left(\frac{\lambda}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k}$$

$$= \frac{\lambda^k}{k!} \cdot \frac{n!}{(n-k)!n^k} \cdot \left(1-\frac{\lambda}{n}\right)^n \cdot \underbrace{\left(1-\frac{\lambda}{n}\right)^{-k}}_{\to 1}$$

$$\to \frac{\lambda^k}{k!} \cdot 1 \cdot e^{-\lambda} \cdot 1 = \boxed{e^{-\lambda}\frac{\lambda^k}{k!}} \quad \text{as } n \to \infty.$$

$$\frac{n!}{(n-k)!\,n^k} = \frac{\overbrace{n(n-1)\times\cdots\times(n-k+1)}^{k \text{ times}}}{\underbrace{n \times n \times \cdots \times n}_{k \text{ times}}} \to \boxed{1} \text{ as } n \to \infty$$

Example

$p = 0.02$

Suppose that the a microchip is defective with probability 0.02. Find the probability that a sample of 100 microchips will contain at most 1 defective microchip.

$n = 100$

$\lambda = np = 2$

$X =$ # of defective mc out of 100 mc's $\sim$ Binomial.

$P(X \leq 1) = p(0) + p(1)$

Example

Suppose that the a microchip is defective with probability $\overset{p=}{\boxed{0.02}}$. Find the probability that a sample of $\underset{\overset{\shortparallel}{n}}{\boxed{100}}$ microchips will contain at most 1 defective microchip.

The number X of defective microchips has a binomial distribution with $n = 100$ and $p = 0.02$. Thus, we obtain

$$p(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$\begin{aligned} P(X \leq 1) &= p(0) + p(1) \\ &= \underline{\binom{100}{0}(0.98)^{100}} + \underline{\binom{100}{1}(0.02)(0.98)^{99}} \approx \boxed{0.403} \end{aligned}$$

Approximately X has a Poisson distribution with $\lambda = (100)(0.02) = \boxed{2}$. We can calculate

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

$$P(X \leq 1) = p(0) + p(1) = \underline{e^{-2}} + \underline{(2)(e^{-2})} \approx \boxed{0.406}$$

Expectation and variance of Poisson distribution.

$X_n \sim$ Binomial with n and $p = \frac{\lambda}{n}$ so that $np = \lambda$

$E[X_n] = np = \lambda$

$X_n \to Y \sim$ Poisson when $n \to \infty$

$Var(X_n) = np(1-p) = \lambda(1 - \frac{\lambda}{n})$

Let X_n, $n = 1, 2, \ldots$, be a sequence of binomial random variables with parameter $(n, \lambda/n)$. Then we regard the limit of the sequence as a Poisson random variable Y, and use the limit to find $E[Y]$ and $\mathrm{Var}(Y)$. Thus, we obtain

$$E[X_n] = \lambda \quad \to \quad E[Y] = \lambda \quad \text{as } n \to \infty;$$

$$\mathrm{Var}(X_n) = \lambda\left(1 - \frac{\lambda}{n}\right) \quad \to \quad \mathrm{Var}(Y) = \lambda \quad \text{as } n \to \infty.$$

Sum of two independent random variables.

Recall: If X and Y are binomial with (n, p) and (m, p) and they are independent then $Z = X+Y$ is binomial with $(n+m, p)$

Theorem

The sum of independent Poisson random variables is a Poisson random variable: If X and Y are independent Poisson random variables with respective parameters λ_1 and λ_2, then $Z = X + Y$ is distributed as the Poisson distribution with parameter $\lambda_1 + \lambda_2$.

Sum of two independent random variables.

Theorem

The sum of independent Poisson random variables is a Poisson random variable: If X and Y are independent Poisson random variables with respective parameters λ_1 and λ_2, then $\underline{Z = X + Y}$ is distributed as the Poisson distribution with parameter $\lambda_1 + \lambda_2$.

$M(t) = e^{\lambda(e^t-1)}$

By the mgf property for independent random variables we obtain

$$M_Z(t) = M_X(t) \times M_Y(t) = e^{\boxed{\lambda_1}(e^t-1)} \times e^{\boxed{\lambda_2}(e^t-1)} = e^{(\lambda_1+\lambda_2)(e^t-1)} = \underline{e^{\lambda(e^t-1)}}$$

where $\lambda = \lambda_1 + \lambda_2$

This uniquely determines the distribution for Z, which is Poisson with parameter $\lambda_1 + \lambda_2$.

This must be a mgf of Poisson.

Alternative approach to the proof.

Idea: $p_n \to 0$ as $k_n, l_n \to \infty$ while $k_n p_n = \lambda_1$ and $l_n p_n = \lambda_2$

↑

You can do it by setting $k_n = \left\lfloor \frac{\lambda_1}{p_n} \right\rfloor$ and $l_n = \left\lfloor \frac{\lambda_2}{p_n} \right\rfloor$ ← $\lfloor \cdot \rfloor$ is floor function (greatest integer)

$\lfloor 11.3 \rfloor = 11$

$\lfloor 295.6 \rfloor = 295$

$k_n / \left(\frac{\lambda_1}{p_n}\right) \to 1$ as $n \to \infty$ $\Leftrightarrow$ $k_n p_n \to \lambda_1$ as $n \to \infty$

Let X_n and Y_n be independent binomial random variables with respective parameter (k_n, p_n) and (l_n, p_n). Then the sum $X_n + Y_n$ of the random variables has the binomial distribution with parameter $(\underbrace{k_n + l_n}_{m_n}, p_n)$. By letting $k_n p_n \to \lambda_1$ and $l_n p_n \to \lambda_2$ with $k_n, l_n \to \infty$ and $p_n \to 0$, the respective limits of X_n and Y_n are Poisson random variables X and Y with respective parameters λ_1 and λ_2. Meanwhile, the limit of $X_n + Y_n$ is the sum of the random variables X and Y, and has a Poisson distribution with parameter $\lambda_1 + \lambda_2$.

$Z_n = X_n + Y_n \sim$ Binomial with $m_n = k_n + l_n$ and $p_n \to 0$ as $n \to \infty$

↓ as $n \to \infty$

$Z = X + Y \sim$ Poisson with $\lambda = \lim_{n\to\infty} m_n p_n = \lim_{n\to\infty} (k_n p_n + l_n p_n) = \lambda_1 + \lambda_2$

Example

Suppose that at an office the average number of calls per hour is 2. (a) If the number X of calls for the first hour has a Poisson distribution, what is the parameter λ? (b) Let Y be the number of calls for three hours. What is the distribution for Y? what is the value for parameter? (c) What is the probability that the first call comes after 3 hours? (d) In general what is the probability that the first call comes after t hours?

rate per unit (an hour)

$E[X] = 2 \leftrightarrow E[X] = \lambda \Rightarrow \lambda = 2$

↑ expectation = average

X_1, X_2, X_3 are number of calls in 1st, 2nd, 3rd hour.

$$Y = \underbrace{\underbrace{X_1 + X_2}_{\text{Poisson with } \lambda = 4} + \underbrace{X_3}_{\text{Poisson with } \lambda = 2}}_{\text{Poisson with } \lambda = 6} \sim \text{Poisson with } \lambda = 6$$

(c) $P(Y = 0) = p(0)$

$= e^{-\lambda} \frac{\lambda^k}{k!}$ with $k = 0$

$= e^{-\lambda} = e^{-6}$

(d) $Y = \underbrace{X_1 + X_2 + \cdots + X_t}_{\text{Poisson with } \lambda = 2t}$

$P(Y=0) = e^{-\lambda} = e^{-2t}$

Example

Suppose that at an office the average number of calls per hour is 2. (a) If the number X of calls for the first hour has a Poisson distribution, what is the parameter λ? (b) Let Y be the number of calls for three hours. What is the distribution for Y? what is the value for parameter? (c) What is the probability that the first call comes after 3 hours? (d) In general what is the probability that the first call comes after t hours?

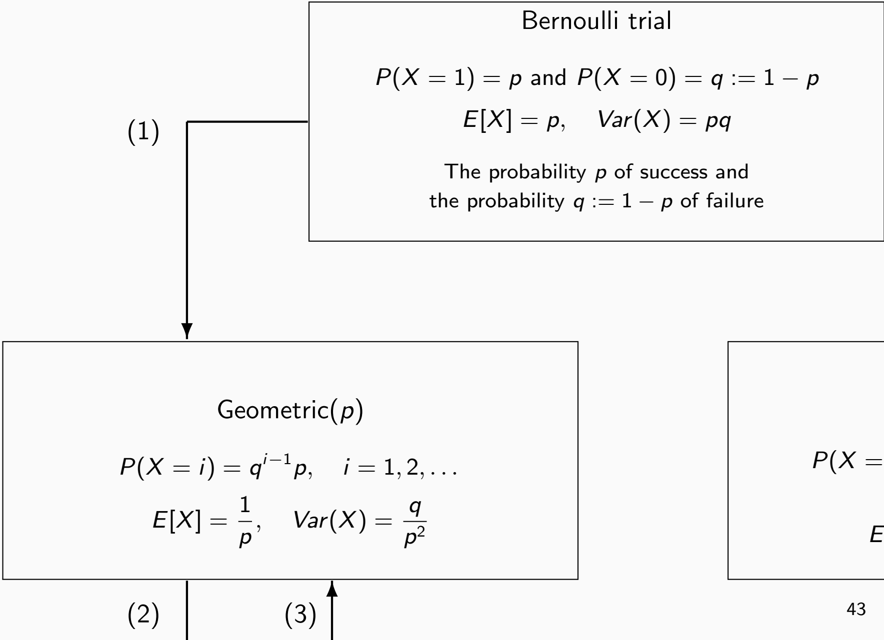
(a) Since $\underline{E[X] = \lambda}$, it must be $\underline{\lambda = 2}$.

(b) Let X_i be the number of calls during the i-th hour. Then $Y = X_1 + X_2 + X_3$ must have a Poisson distribution with parameter $\lambda = (2)(3) = 6$.

(c) The probability is obtained as $P(Y = 0) = e^{-6} \approx 0.002$.

(d) Let Y be the number of calls for t hours. Then the probability is obtained as $P(Y = 0) = e^{-2t}$.

Summary diagram.



Bernoulli trial

$$P(X = 1) = p \text{ and } P(X = 0) = q := 1 - p$$

$$E[X] = p, \quad Var(X) = pq$$

The probability p of success and
the probability $q := 1 - p$ of failure

(1)

Geometric(p)

$$P(X = i) = q^{i-1}p, \quad i = 1, 2, \ldots$$

$$E[X] = \frac{1}{p}, \quad Var(X) = \frac{q}{p^2}$$

$P(X =$

E

(2)

(3)

Relationship among random variables.

(1) X is the number of Bernoulli trials until the *first* success occurs.

(2) X is the number of Bernoulli trials until the rth success occurs. If $X_1, \cdots, X_r$ are independent geometric random variables with parameter p, then $X = \sum_{k=1}^{r} X_k$ becomes a negative binomial random variable with parameter (r, p).

(3) $r = 1$.

(4) X is the number of successes in n Bernoulli trials. If $X_1, \cdots, X_n$ are independent Bernoulli trials, then $X = \sum_{k=1}^{n} X_k$ becomes a binomial random variable with (n, p).

(5) If X_1 and X_2 are independent binomial random variables with respective parameters (n_1, p) and (n_2, p), then $X = X_1 + X_2$ is *also* a binomial random variable with $(n_1 + n_2, p)$.

Relationship among random variables, continued.

(6) Poisson approximation is used for binomial distribution by letting $n \to \infty$ and $p \to 0$ while $\lambda = np$.

(7) If X_1 and X_2 are independent Poisson random variables with respective parameters λ_1 and λ_2, then $X = X_1 + X_2$ is *also* a Poisson random variable with $\lambda_1 + \lambda_2$.

Homework on Problems 1-4 and

Optional problems 14 and 15.

Quiz on Oct. 24

Optional problems are due Oct 27.

Problem 1

A study shows that 40% of college students binge drink. Let X be the number of students who binge drink out of sample size $n = 12$.

1. *Find the mean and standard deviation of X.*
2. *Do you agree that the probability that X is 5 or less is higher than 50%? Justify your answer.*

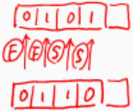
$X \sim$ Binomial with $(n = 12, p = 0.4)$

(b) $P(X \leq 5) = p(0) + p(1) + \cdots + p(5) \leftarrow p(k) = \underline{\binom{12}{k}(0.4)^k(0.6)^{12-k}}$

Problem

In a lot of 20 light bulbs, there are 9 bad bulbs. Let X be the number of defective bulbs found in the inspection. Find the frequency function $p(k)$, and identify the range for k in the following inspection procedures.

1. *An inspector inspects 5 bulbs selected at random and without replacement.*
2. *An inspector inspects 15 bulbs selected at random and without replacement.*



$X = \#$ of error bits $= 2$ ← success probability $p=0.05$

$= x_1 + x_2 + x_3 + x_4$ ⇝ Binomial with $n=4$, $p=0.05$

$\underbrace{x_1 + x_2 + x_3 + x_4}_{\text{They are independent}}$

Problem

Appending three extra bits to a 4-bit word in a particular way (a Hamming code) allows detection and correction of up to one error in any of the bits. (a) If each bit has probability .05 of being changed during communication, and the bits are changed independently of each other, what is the probability that the word is correctly received (that is, 0 or 1 bit is in error)? (b) How does this probability compare to the probability that the word will be transmitted correctly with no check bits, in which case all four bits would have to be transmitted correctly for the word to be correct?

$P(X=k) = \binom{n}{k} p^k (1-p)^{n-k}$

(a) $P(X \le 1) = P(X=0) + P(X=1) = \binom{4}{0}(0.05)^0(0.95)^4 + \binom{4}{1}(0.05)^1(0.95)^3$

(b) $P(X=0) = \binom{4}{0}(0.05)^0(0.95)^4$

Problem

Which is more likely: 9 heads in 10 tosses of a fair coin or 18 heads in 20 tosses?

$p(1) = P(X=1) = p_1$

$p(2) = P(X=2) = (1-p_1)p_2$

$p(3) = P(X=3) = (1-p_1)(1-p_2)p_1$

$p(4) = P(X=4) = (1-p_1)^2(1-p_2)p_2$

Problem 5

Two boys play basketball in the following way. They take turns shooting and stop when a basket is made. Player A goes first and has probability p_1 of making a basket on any throw. Player B goes, who shoots second, has probability p_2 of making a basket. The outcomes of the successive trials are assumed to be independent.

1. *Find the frequency function for the total* $\underbrace{\text{number of attempts}}_{X}$.
2. *What is the probability that player A wins?*

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

$p(2k-1) = P(A \text{ wins at } k\text{-th attempt}) = (1-p_1)^{k-1}(1-p_2)^{k-1}p_1$

$p(2k) = P(B \text{ wins at } k\text{-th attempt}) =$

(a) $p(k) = P(X = k) = P(Y-1=k) = P(Y = k+1)$

Geometric

$X \rightarrow Y = X+1 =$ # of trials until 1st success

Problem 6

Suppose that in a sequence of independent Bernoulli trials each with probability of success p, the number of failure up to the first success is counted. (a) What is the frequency function for this random variable? (b) Find the frequency function for the number of failures up to the rth success.

Hint: In (b) let X be the number of failures up to the rth success, and let Y be the number of trials until the r-th success. What is the relation between the random variables X and Y?

$Y = X + r =$ # of trials until r-th success.

Negative Binomial

(b) $p(k) = P(X = k) = P(Y = k+r)$

Problem

Find an expression for the cumulative distribution function of a geometric random variable.

Problem

In a sequence of independent trials with probability p of success, what is the probability that there are exactly r successes before the kth failure?

Problem

The probability that a man in a certain age group dies in the next four years is $p = 0.05$*. Suppose that we observe* 20 *men.*

1. *Find the probability that two of them or fewer die in the next four years.*
2. *Find an approximation by using a Poisson distribution.*

# of successes out of n trials = $X \sim$ Binomial with (n, p) $Y \sim$ Poisson with λ

as $n \to \infty$ If n is large

$$\lim_{n\to\infty} P(X=k) = P(Y=k) \rightarrow P(X=k) \approx P(Y=k) \text{ with } \boxed{np=\lambda}$$

while $np = \lambda$

Problem 10

The probability of being dealt a royal straight flush (ace, king, queen, jack, and ten of the same suit) in poker is about $\overset{p}{1.5 \times 10^{-6}}$*. Suppose that an avid poker player sees* 100 *hands a week,* 52 *weeks a year, for* 20 *years.*

$n = (100)(52)(20)$ → $\boxed{\lambda = 0.156}$

1. *What is the probability that she never sees a royal straight flush dealt?* $P(X=0) \approx P(Y=0) = p(0) = e^{-\lambda}\frac{\lambda^0}{0!} = e^{-0.156}$
2. *What is the probability that she sees at least two royal straight flushes dealt?*

$$P(X \geq 2) = 1 - P(X \leq 1) \approx 1 - P(Y \leq 1) = 1 - p(0) - p(1)$$

$$= 1 - e^{-\lambda}\frac{\lambda^0}{0!} - e^{-\lambda}\frac{\lambda^1}{1!} = 1 - e^{-\lambda} - \lambda e^{-\lambda}$$

$$p(k) = \binom{n}{k} p^k (1-p)^{n-k}$$

Problem

Professor Rice was told that he has only 1 chance in 10,000 of being trapped in a much-maligned elevator in the mathematics building. Assume that the outcomes on all the days are mutually independent. If he goes to work 5 days a week, 52 weeks a year, for 10 years and always rides the elevator up to his office when he first arrives.

1. *What is the probability that he will never be trapped?*
2. *What is the probability that he will be trapped once?*
3. *What is the probability that he will be trapped twice?*

Problem

Suppose that a rare disease has an incidence of 1 *in* 1000. *and that members of the population are affected independently. Let* X *be the number of cases in a population of* $100,000$.

1. *What is the average number of cases?*
2. *Find the frequency function.*

X_i = # of suicides in ith month

$P(X_1=2) = p(2) = e^{-\lambda}\frac{\lambda^2}{2!}$

Problem

Suppose that in a city the number of suicides can be approximated by a Poisson random variable with $\lambda = 0.33$ *per month.*

1. *What is the probability of two suicides in January?*
2. *What is the distribution for the number of suicides per year?*
3. *What is the average number of suicides in one year?*
4. *What is the probability of two suicides in one year?*

Y = # of suicides in a year $= X_1 + \cdots + X_{12}$

$\sim$ Poisson with parameter $\lambda = (12)(0.33)$

If X_1 and X_2 are independent Poisson with λ_1 and λ_2

$Y = X_1 + X_2 \sim$ Poisson with parameter $\lambda_1 + \lambda_2$

Optional problems

Problem

If n men throw their hats into a pile and each man takes a hat at random, what is the expected number of matches?
Hint: Suppose Y is the number of matches. How can you express the random variable Y by using indicator random variables $X_1, \ldots, X_n$?
What is the event for the indicator random variable X_i?

Problem

Suppose that n enemy aircraft are shot at simultaneously by m gunners, that each gunner selects an aircraft to shoot at independently of the other gunners, and that each gunner hits the selected aircraft with probability p. Find the expected number of aircraft hit by the gunners. Hint: For each $i = 1, \ldots, n$, *let* X_i *be the indicator random variable of the event that the i-th aircraft was hit. Explain how you can verify* $P(X_i = 0) = \left(1 - \frac{p}{n}\right)^m$ *by yourself, and use it to find the expectation of interest.*

Problem

(Banach Match Problem) A pipe smoker carries one box of matches in his left pocket and one in his right. Initially, each box contains n matches. If he needs a match, the smoker is equally likely to choose either pocket. What is the frequency function for the number of matches in the other box when he first discovers that one box is empty?

Hint: Let A_k *denote the event that the pipe smoker first discovers that the right-hand box is empty and that there are k matches in the left-hand box at the time, and similarly let* B_k *denote the event that he first discovers that the left-hand box is empty and that there are k matches in the right-hand box at the time. Then use* A_k*'s and* B_k*'s to express the frequency function of interest.*

Problem 17

Let X be a negative binomial random variable with parameter (r, p).

\# of trials until 2nd successes

1. *When $r = 2$ and $p = 0.2$, find the probability $P(X \leq 4)$.*

   \# of successes
2. *Let Y be a binomial random variable with parameter (n, p). When $n = 4$ and $p = 0.2$, find the probability $P(Y \geq 2)$.* $= P(X \leq 4)$
3. *By comparing the results above, what can you find about the relationship between $P(X \leq 4)$ and $P(Y \geq 2)$? Generalize the relationship to the one between $P(X \leq n)$ and $P(Y \geq 2)$, and find the probability $P(X \leq 10)$ when $r = 2$ and $p = 0.2$.*
4. *By using the Poisson approximation, find the probability $P(X \leq 1000)$ when $r = 2$ and $p = 0.001$.*

Answers to exercises

Problem 1.

1. The mean is $np = (12)(0.4) = 4.8$, and the standard deviation is $\sqrt{(12)(0.4)(0.6)} \approx 1.7$.
2. Yes, because $P(X \leq 5) \approx 0.665$.

Problem 2.

X has a hypergeometric distribution with $N = 20$ and $m = 9$.

1. Here we choose $r = 5$. Thus,

$$p(k) = \frac{\binom{9}{k}\binom{11}{5-k}}{\binom{20}{5}}$$

   where k takes a value in the range of $0 \leq k \leq 5$.

2. Here we choose $r = 15$. Thus,

$$p(k) = \frac{\binom{9}{k}\binom{11}{15-k}}{\binom{20}{15}}$$

   where k takes a value in the range of $4 \leq k \leq 9$.

Problem 3.

The number X of error bits has a binomial distribution with $n = 4$ and $p = 0.05$.

(a) $P(X \leq 1) = \binom{4}{0}(0.95)^4 + \binom{4}{1}(0.05)(0.95)^3 \approx 0.986.$

(b) $P(X = 0) = \binom{4}{0}(0.95)^4 \approx 0.815.$

Problem 4.

“9 heads in 10 tosses” $[\binom{10}{9}(1/2)^{10} \approx 9.77 \times 10^{-3}]$ is more likely than “18 heads in 20 tosses” $[\binom{20}{18}(1/2)^{20} \approx 1.81 \times 10^{-4}]$.

Problem 5.

Let X be the total number of attempts.

(a) $P(X = n) = \begin{cases} (1-p_1)^{k-1}(1-p_2)^{k-1}p_1 & \text{if } n = 2k-1; \\ (1-p_1)^k(1-p_2)^{k-1}p_2 & \text{if } n = 2k. \end{cases}$

(b) $P(\{\text{Player A wins}\}) = \sum_{k=1}^{\infty} P(X = 2k-1) =$

$p_1 \sum_{k=1}^{\infty} [(1-p_1)(1-p_2)]^{k-1} = \frac{p_1}{p_1 + p_2 - p_1 p_2}$

Problem 6.

Let X be the number of failures up to the rth success, and let Y be the number of trials until the r-th success. Then Y has a negative binomial distribution, and has the relation $X = Y - r$.

(a) Here $r = 1$ (that is, Y has a geometric distribution), and $P(X = k) = P(Y = k + 1) = (1 - p)^k p$.

(b) $P(X = k) = P(Y = k + r) = \binom{k + r - 1}{r - 1} p^r (1 - p)^k$.

Problem 7.

Let X be a geometric random variable.

$$P(X \leq k) = \sum_{i=1}^{k} p(1-p)^{i-1} = 1 - (1-p)^k$$

Problem 8.

Let X be the number of trials up to the k-th failure. Then X has a negative binomial distribution with "success probability" $(1-p)$.

$$\begin{aligned} &P(\{\text{Exactly } r \text{ successes before the } k\text{-th failure}\}) \\ &= P(X = r+k) = \binom{k+r-1}{k-1}(1-p)^k p^r \end{aligned}$$

Problem 9.

1. The number X of men dying in the next four years has a binomial distribution with $n = 20$ and $p = 0.05$.

$$\begin{aligned} &P(X \leq 2) = p(0) + p(1) + p(2) \\ &= \binom{20}{0}(0.95)^{20} + \binom{20}{1}(0.05)(0.95)^{19} + \binom{20}{2}(0.05)^2(0.95)^{18} \\ &\approx 0.925 \end{aligned}$$

2. The random variable X has approximately a Poisson distribution with $\lambda = (0.05)(20) = 1$.

$$P(X \leq 2) = e^{-1} + e^{-1} + \frac{1}{2}e^{-1} = \frac{5}{2}e^{-1} \approx 0.920$$

Problem 10.

1. The number X of royal straight flushes has a Poisson distribution with

   $$\lambda = 1.5 \times 10^{-6} \times (100 \times 52 \times 20) = 0.156$$

   Thus, we obtain

   $$P(X = 0) = p(0) = e^{-0.156} \approx 0.856$$

2. $P(X \geq 2) = 1 - p(0) - p(1) = 1 - e^{-0.156} - (0.156)(e^{-0.156}) = 1 - (1.156)(e^{-0.156}) \approx 0.011$

Problem 11.

1. The number X of misfortunes for Professor Rice being trapped has a Poisson distribution with

   $$\lambda = 10^{-4} \times (5 \times 52 \times 10) = 0.26$$

   Thus, we obtain

   $$P(X = 0) = p(0) = e^{-0.26} \approx 0.771$$

2. $P(X = 1) = p(1) = (0.26)(e^{-0.26}) \approx 0.20$
3. $P(X = 2) = p(2) = \frac{(0.26)^2}{2}(e^{-0.26}) \approx 0.026$

Problem 12.

1. The number X of cases has a Poisson distribution with

   $$\lambda = 10^{-3} \times 100,000 = 100$$

   Thus, we obtain $E[X] = \lambda = 100$.
2. $p(k) = (e^{-100})\dfrac{100^k}{k!}$ for $k = 0, 1, \ldots.$

Problem 13.

1. Let X_1 be the number of suicides in January. Then we have

$$P(X_1 = 2) = \frac{(0.33)^2}{2}(e^{-0.33}) \approx 0.039$$

2. Let X_1,...,X_{12} be the number of suicides in i-th month. Then

$$Y = X_1 + X_2 + \cdots + X_{12}$$

   represents the number of suicides in one year, and it has a poisson distribution.

3. $E[Y] = E[X_1] + E[X_2] + \cdots + E[X_{12}] = (12)(0.33) = 3.96.$
4. Since $\lambda = E[Y] = 3.96$ is the parameter value for the Poisson random variable Y, we obtain

$$P(Y = 2) = \frac{(3.96)^2}{2}(e^{-3.96}) \approx 0.149$$