S&DS 241 Lecture 8

Distributions related to independent trials: Bern, Bin, Geo

B-H: 3.3, 4.3, math appendix A.8

(Example 1.5.1-1.5.3 for binomial identities)

Bernoulli distribution

- A random variable X has a Bernoulli distribution with parameter p, denoted by $X \sim \text{Bern}(p)$, if $P(X = 1) = p$ and $P(X = 0) = 1 - p$.
- Interpretation: Bernoulli trial

$$X = \begin{cases} 1 & \text{success} \\ 0 & \text{fail} \end{cases}$$

Bernoulli distribution

- A random variable X has a Bernoulli distribution with parameter p, denoted by $X \sim \text{Bern}(p)$, if $P(X = 1) = p$ and $P(X = 0) = 1 - p$.
- Interpretation: Bernoulli trial

$$X = \begin{cases} 1 & \text{success} \\ 0 & \text{fail} \end{cases}$$

J. Bernoulli

Bernoulli distribution

- A random variable X has a Bernoulli distribution with parameter p, denoted by $X \sim \text{Bern}(p)$, if $P(X = 1) = p$ and $P(X = 0) = 1 - p$.
- Interpretation: Bernoulli trial

$$X = \begin{cases} 1 & \text{success} \\ 0 & \text{fail} \end{cases}$$

J. Bernoulli

Bernie

Binomial distribution

Binomial distribution arises in independent trials (repeated experiments):

- Perform n independent Bernoulli trials, each of which succeeds with probability p
- Let $X =$ the number of successes
- We say X has a binomial distribution with parameter n and p, denoted by $X \sim \text{Bin}(n, p)$,
- Clearly, $\text{Bin}(1, p)$ is just $\text{Bern}(p)$

Binomial = sum of independent Bernoullis

- Define the indicator random variables

$$X_i = \begin{cases} 1 & i^{\text{th}} \text{ trial succeeds} \\ 0 & i^{\text{th}} \text{ trial fails} \end{cases}$$

- $X_1, \ldots, X_n$ are independent and identically distributed (iid) according to Bern(p)
- Then

$$X = \sum_{i=1}^{n} X_i$$

is the total number of successes

Binomial PMF

Let $X \sim \text{Bin}(n, p)$. How to find its PMF?

Binomial PMF

Let $X \sim \text{Bin}(n, p)$. How to find its PMF?

- X takes values in $\{0, 1, \dots, n\}$

Binomial PMF

Let $X \sim \text{Bin}(n, p)$. How to find its PMF?

- X takes values in $\{0, 1, \ldots, n\}$
- $P(X = 0) = (1 - p)^n, P(X = n) = p^n$

Binomial PMF

Let $X \sim \text{Bin}(n, p)$. How to find its PMF?

- X takes values in $\{0, 1, \ldots, n\}$
- $P(X = 0) = (1 - p)^n, P(X = n) = p^n$
- $P(X = 1) = np(1 - p)^{n-1}$

Binomial PMF

Let $X \sim \text{Bin}(n, p)$. Then the PMF of X is

$$P\left(X = k\right) = \binom{n}{k} p^k \left(1 - p\right)^{n-k}, \quad k = 0, 1, \ldots, n$$

Binomial PMF

Let $X \sim \text{Bin}(n, p)$. Then the PMF of X is

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

Proof:

- $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ = number of ways to succeed k out of n trials
- p^k = probability of k successes
- $(1-p)^{n-k}$ = probability of $n-k$ failures

Normalization

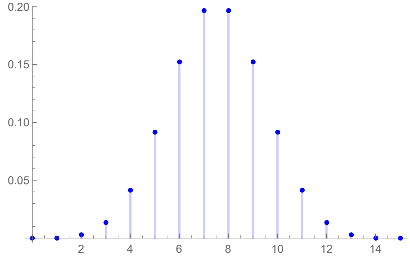
- Recall binomial expansion:

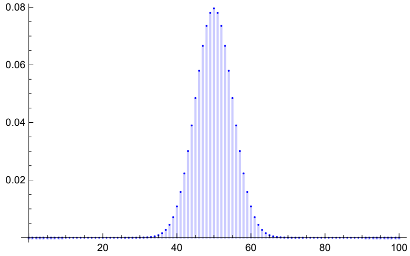
$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k}$$

- Substituting $x = p$ and $y = 1 - p$:

$$\sum_{k=0}^{n} \binom{n}{k} p^k \left(1-p\right)^{n-k} = \left(p + \left(1-p\right)\right)^n = 1$$

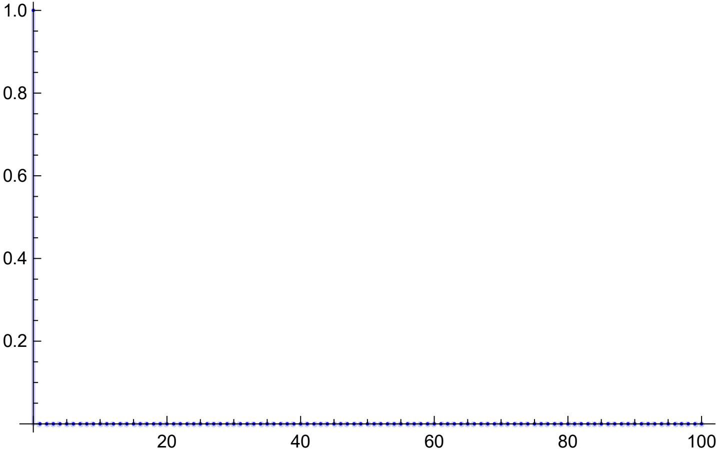
Binomial PMF: Bin$(15, 1/2)$ vs Bin$(100, 1/2)$





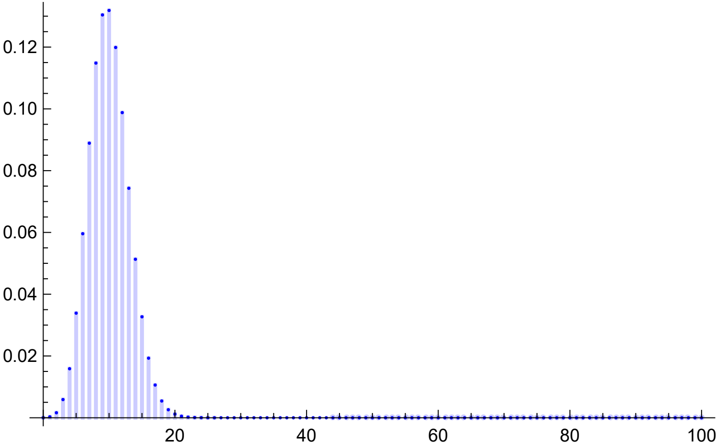
$$X \sim \mathsf{Bin}(n, 1/2), \quad P\left(X = k\right) = \binom{n}{k} \frac{1}{2^n}$$

Binomial PMF



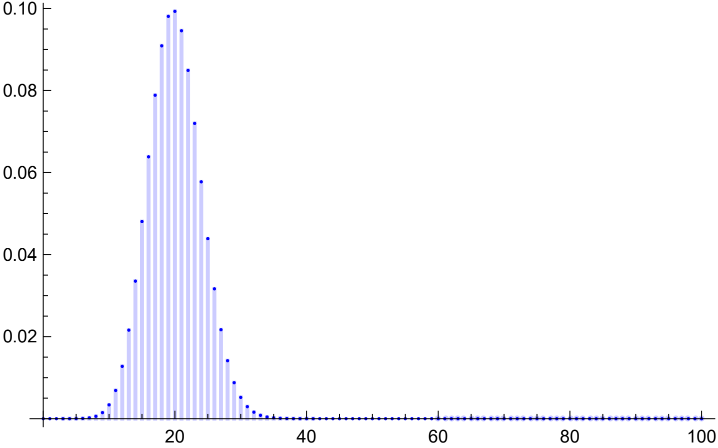
$\text{Bin}(100, 0/10)$

Binomial PMF



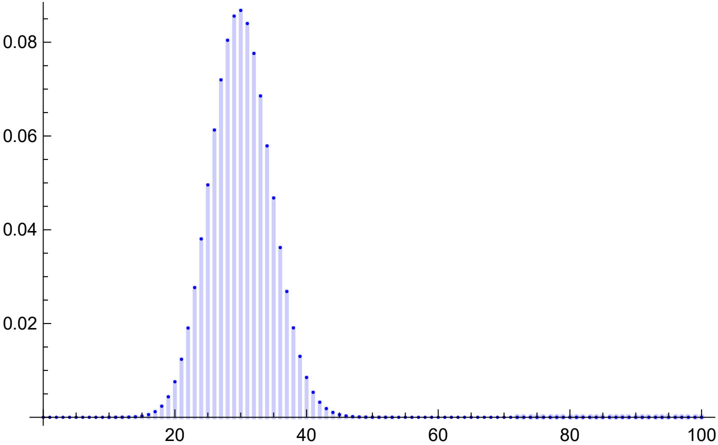
$\text{Bin}(100, 1/10)$

Binomial PMF



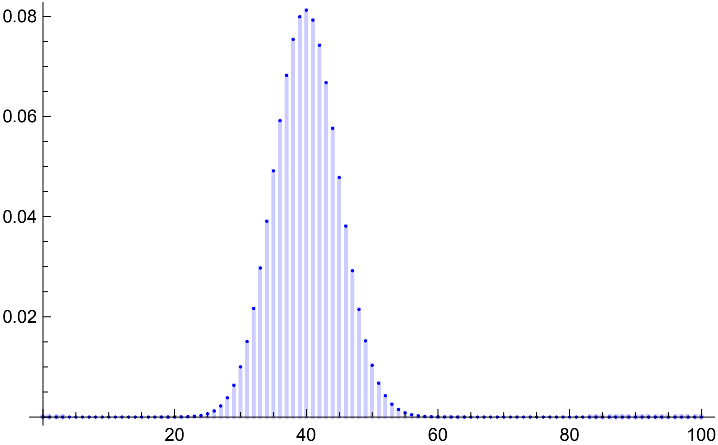
$\mathsf{Bin}(100, 2/10)$

Binomial PMF



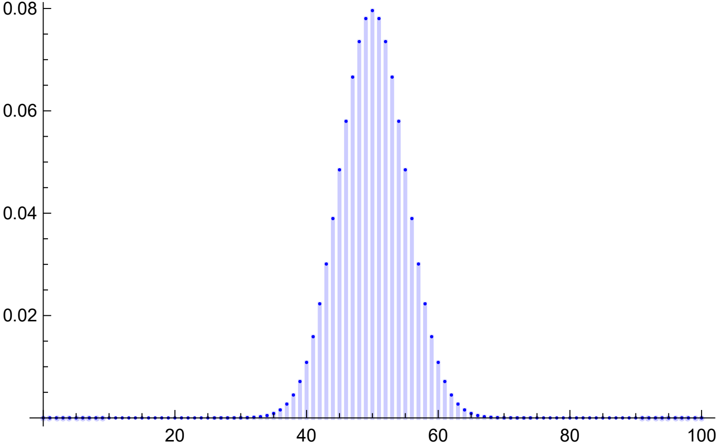
$\text{Bin}(100, 3/10)$

Binomial PMF



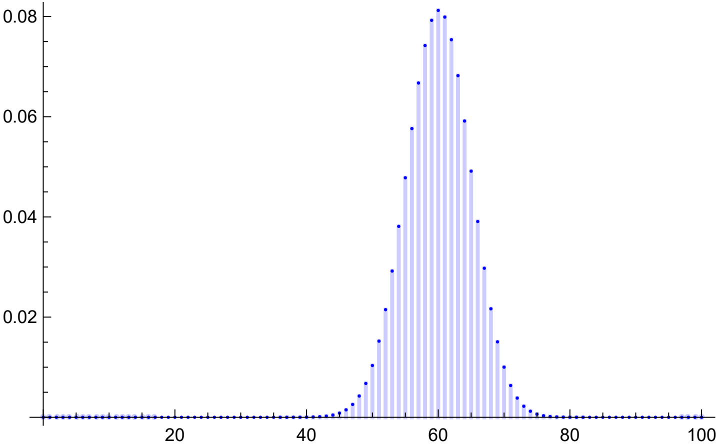
Bin$(100, 4/10)$

Binomial PMF



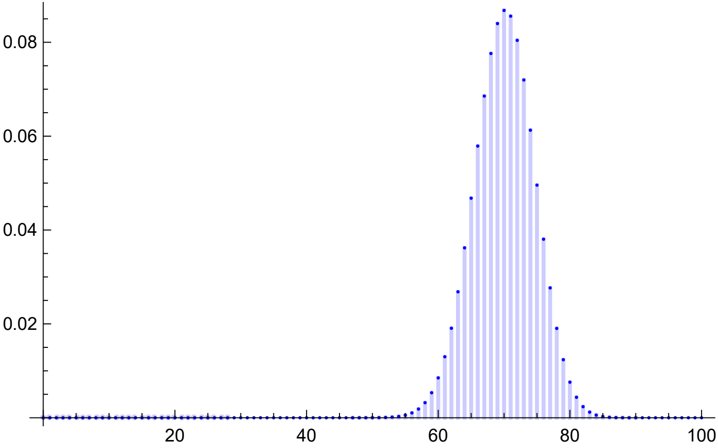
$\text{Bin}(100, 5/10)$

Binomial PMF



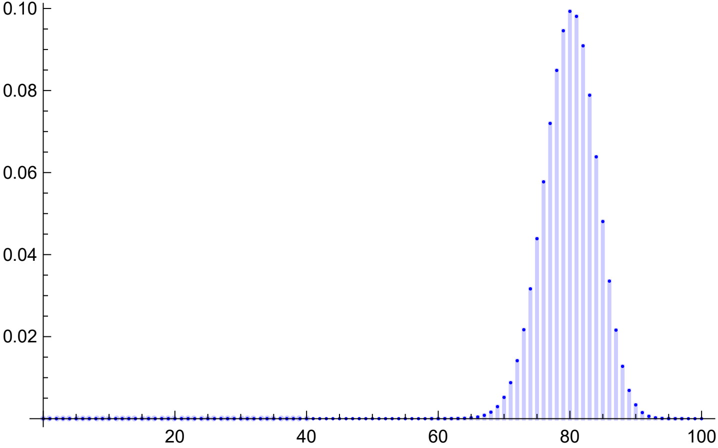
$\text{Bin}(100, 6/10)$

Binomial PMF



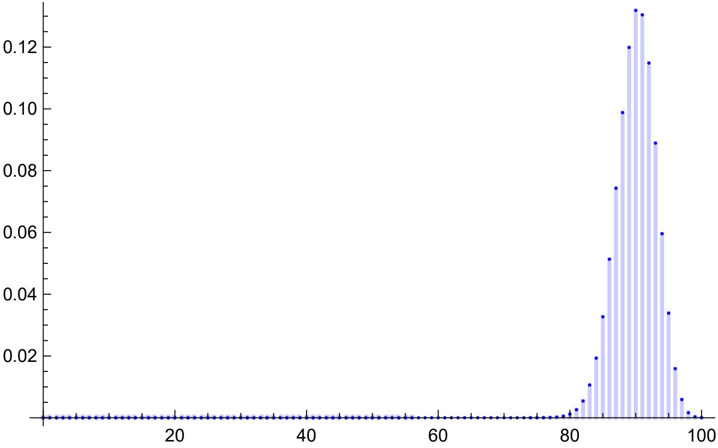
$\text{Bin}(100, 7/10)$

Binomial PMF



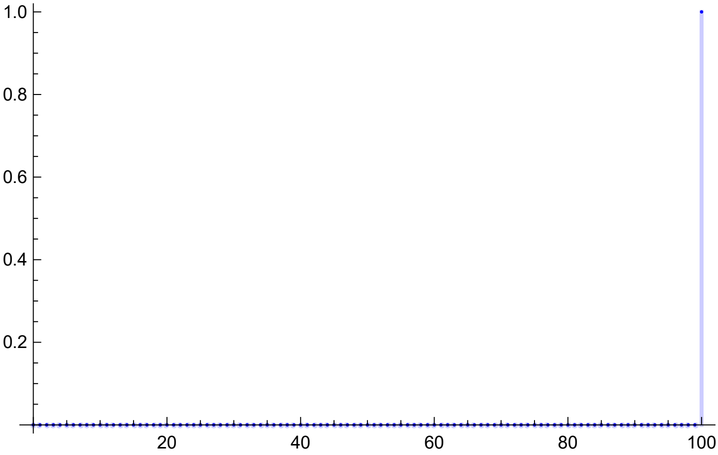
$\text{Bin}(100, 8/10)$

Binomial PMF



$\text{Bin}(100, 9/10)$

Binomial PMF



$\text{Bin}(100, 10/10)$

Mean of Bin(n, p)

$X \sim \mathsf{Bin}(n, p)$

$$E(X) = np$$

Mean of Bin(n, p)

$X \sim \text{Bin}(n, p)$

$$E(X) = np$$

Proof 1 Linearity of expecatation

$$E(X) = E(X_1 + \ldots + X_n) = E(\underbrace{X_1}_{\text{Bern}(p)}) + \ldots + E(\underbrace{X_n}_{\text{Bern}(p)}) = np$$

Mean of Bin(n, p)

$X \sim \text{Bin}(n, p)$

$$E(X) = np$$

Proof 1 Linearity of expecatation

$$E(X) = E(X_1 + \ldots + X_n) = E(\underbrace{X_1}_{\text{Bern}(p)}) + \ldots + E(\underbrace{X_n}_{\text{Bern}(p)}) = np$$

Proof 2 Direct calculation via PMF

$$E(X) = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k} = np \underbrace{\sum_{\ell=0}^{n-1} \binom{n-1}{\ell} p^\ell (1-p)^{n-1-\ell}}_{=1}$$

where we applied $n\binom{n}{k} = k\binom{n-1}{k-1}$ [B-H Example 1.5.2]

Mean of Bin(n, p)

$X \sim \text{Bin}(n, p)$

$$E(X) = np$$

Proof 1 Linearity of expecatation

$$E(X) = E(X_1 + \ldots + X_n) = E(\underbrace{X_1}_{\text{Bern}(p)}) + \ldots + E(\underbrace{X_n}_{\text{Bern}(p)}) = np$$

Proof 2 Direct calculation via PMF

$$E(X) = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k} = np \underbrace{\sum_{\ell=0}^{n-1} \binom{n-1}{\ell} p^\ell (1-p)^{n-1-\ell}}_{=1}$$

where we applied $n\binom{n}{k} = k\binom{n-1}{k-1}$ [B-H Example 1.5.2]
Mode (most likely value) of Bin(n, p) is $\lfloor (n+1)p \rfloor$: (Exercise)

Properties of binomial distributions

Let $X \sim \mathsf{Bin}(n, p)$.

- What is the distribution of $n - X$?

Properties of binomial distributions

Let $X \sim \mathsf{Bin}(n, p)$.

- What is the distribution of $n - X$?

$$\mathsf{Bin}(n, 1-p)$$

  as number of failures in n trials

Properties of binomial distributions

Let $X \sim \mathsf{Bin}(n, p)$.

- What is the distribution of $n - X$?

$$\mathsf{Bin}(n, 1 - p)$$

  as number of failures in n trials

- Let $Y \sim \mathsf{Bin}(m, p)$ be independent of X. What is the distribution of $X + Y$?

Properties of binomial distributions

Let $X \sim \text{Bin}(n, p)$.

- What is the distribution of $n - X$?

$$\text{Bin}(n, 1 - p)$$

  as number of failures in n trials

- Let $Y \sim \text{Bin}(m, p)$ be independent of X. What is the distribution of $X + Y$?

$$\text{Bin}(m + n, p)$$

  as number of successes in $m + n$ trials

Properties of binomial distributions

Let $X \sim \text{Bin}(n, p)$.

- What is the distribution of $n - X$?

  $$\text{Bin}(n, 1-p)$$

  as number of failures in n trials
- Let $Y \sim \text{Bin}(m, p)$ be independent of X. What is the distribution of $X + Y$?

  $$\text{Bin}(m+n, p)$$

  as number of successes in $m + n$ trials

As a warm-up exercise, let's verify these by PMF.

Direct verification

Let $X \sim \mathsf{Bin}(n, p)$. Let $Z = n - X$, wihch is a function of X. Then

Direct verification

Let $X \sim \text{Bin}(n, p)$. Let $Z = n - X$, wihch is a function of X. Then

$$\begin{aligned} P(Z = k) &= P(X = n - k) \\ &= \binom{n}{n-k} p^{n-k}(1-p)^k \\ &= \binom{n}{k}(1-p)^k p^{n-k} \qquad \text{[B-H Example 1.5.1]} \end{aligned}$$

which is the PMF of $\text{Bin}(n, 1-p)$ evaluated at k

Direct verification

Let $X \sim \mathsf{Bin}(n, p)$ and $Y \sim \mathsf{Bin}(m, p)$ be independent. Let $S = X + Y$.

Direct verification

Let $X \sim \mathsf{Bin}(n, p)$ and $Y \sim \mathsf{Bin}(m, p)$ be independent. Let $S = X + Y$.

$$P(S = k) = \sum_{i=0}^{k} P(X = i) P(Y = k - i)$$

Direct verification

Let $X \sim \text{Bin}(n, p)$ and $Y \sim \text{Bin}(m, p)$ be independent. Let $S = X + Y$.

$$
\begin{aligned}
P(S = k) &= \sum_{i=0}^{k} P(X = i) P(Y = k - i) \\
&= \sum_{i=0}^{k} \binom{n}{i} p^i (1-p)^{n-i} \binom{m}{k-i} p^{k-i} (1-p)^{m-k+i}
\end{aligned}
$$

Direct verification

Let $X \sim \text{Bin}(n, p)$ and $Y \sim \text{Bin}(m, p)$ be independent. Let $S = X + Y$.

$$
\begin{aligned}
P(S = k) &= \sum_{i=0}^{k} P(X = i) P(Y = k - i) \\
&= \sum_{i=0}^{k} \binom{n}{i} p^i (1-p)^{n-i} \binom{m}{k-i} p^{k-i} (1-p)^{m-k+i} \\
&= p^k (1-p)^{m+n-k} \underbrace{\sum_{i=0}^{k} \binom{n}{i} \binom{m}{k-i}}_{=\binom{m+n}{k} \text{ [B-H Example 1.5.3]}}
\end{aligned}
$$

which is the PMF of $\text{Bin}(m + n, p)$ evaluated at k

Example: Best of five

Alice plays against Bob in a best-of-five match, and wins each game with probability $1/2$ independently. Let Y be the total number of games played. Find PMF of Y.

Example: Best of five

Alice plays against Bob in a best-of-five match, and wins each game with probability $1/2$ independently. Let Y be the total number of games played. Find PMF of Y.

- Y can be $3, 4$ or 5

Example: Best of five

Alice plays against Bob in a best-of-five match, and wins each game with probability $1/2$ independently. Let Y be the total number of games played. Find PMF of Y.

- Y can be $3, 4$ or 5
- Denote the score by A:B

Example: Best of five

Alice plays against Bob in a best-of-five match, and wins each game with probability $1/2$ independently. Let Y be the total number of games played. Find PMF of Y.

- Y can be $3, 4$ or 5
- Denote the score by A:B
- $P(Y = 3) = P(3 : 0) + P(0 : 3) = \frac{1}{8} + \frac{1}{8} = \frac{1}{4}$

Example: Best of five

$$
\begin{aligned}
P(Y=4) &= P(3:1) + P(1:3) \\
&= 2P(3:1) \\
&= 2P(\text{A wins 2 out of the first 3 games}) \times P(\text{A wins 4th game}) \\
&= 2\binom{3}{2}\frac{1}{2^3} \times \frac{1}{2} \\
&= \frac{3}{8}
\end{aligned}
$$

since the number of games won by A in the first 3 games $\sim$ $\text{Bin}(3, 1/2)$

Example: Best of five

$$\begin{aligned}
P(Y=5) &= P(3:2) + P(2:3) \\
&= 2P(3:2) \\
&= 2P(\text{A wins 2 out of the first 4 games}) \times P(\text{A wins 5th game}) \\
&= 2\binom{4}{2}\frac{1}{2^4} \times \frac{1}{2} \\
&= \frac{3}{8}
\end{aligned}$$

since the number of games won by A in the first 4 games $\sim$ Bin$(4, 1/2)$

Geometric distribution: time till first success

- Perform independent Bernoulli trials with success p.
- Define the random variable

$$L = \text{number of failures till the first success}$$

- We say L follows a geometric distribution with parameter p, denoted by $L \sim \mathsf{Geom}(p)$

PMF of Geom(p)

- L takes values in $\{0, 1, 2, 3, \ldots\}$, the set of non-negative integers (countably infinite)

PMF of Geom(p)

- L takes values in $\{0, 1, 2, 3, \ldots\}$, the set of non-negative integers (countably infinite)
- PMF: for each $k \geq 0$,

$$\begin{aligned} P(L = k) &= P(\text{fail first } k \text{ attempts, succeed in the } (k+1)^{\text{th}}) \\ &= (1-p)^k p \end{aligned}$$

which decays geometrically as k increases

PMF of Geom(p)

- L takes values in $\{0, 1, 2, 3, \ldots\}$, the set of non-negative integers (countably infinite)
- PMF: for each $k \geq 0$,

$$\begin{aligned} P(L = k) &= P(\text{fail first } k \text{ attempts, succeed in the } (k+1)^{\text{th}}) \\ &= (1-p)^k p \end{aligned}$$

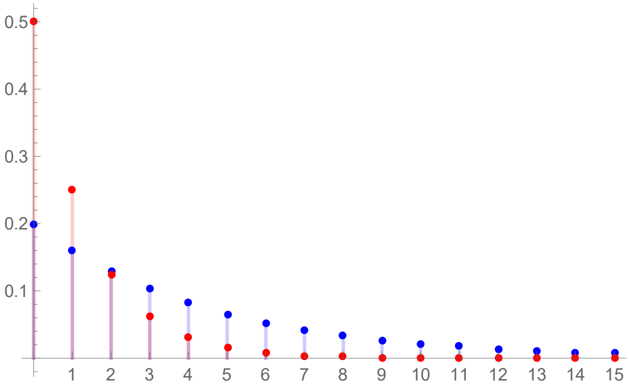
  which decays geometrically as k increases
- Sanity check:

$$\sum_{k \geq 0} P(L = k) = \sum_{k \geq 0} (1-p)^k p = 1$$

  Recall geometric series (B-H Math Appendix A.8.2):

$$\sum_{k \geq k_0} \alpha^k = \frac{\alpha^{k_0}}{1-\alpha} = \boxed{\frac{\text{first term}}{1 - \text{ratio}}}$$

$\color{red}{\text{Geom}(0.5)}$ vs $\color{blue}{\text{Geom}(0.2)}$



Mean of Geom(p)

$L \sim \mathsf{Geom}(p)$

$$E(L) = \frac{1-p}{p}$$

Interpretation: average number of attempts to reach first success is inversely proportional to the success probability

Mean of Geom(p)

$L \sim \mathsf{Geom}(p)$

$$E(L) = \frac{1-p}{p}$$

Interpretation: average number of attempts to reach first success is inversely proportional to the success probability

B-H Example 4.3.6.

$$E(L) = \sum_{k\geq 0} kP(L=k) = \sum_{k\geq 1} k(1-p)^k p \overset{(*)}{=} \frac{1}{p^2} \times (1-p)p = \frac{1-p}{p}$$

Mean of Geom(p)

$L \sim \text{Geom}(p)$

$$E(L) = \frac{1-p}{p}$$

Interpretation: average number of attempts to reach first success is inversely proportional to the success probability

B-H Example 4.3.6.

$$E(L) = \sum_{k\geq 0} kP(L=k) = \sum_{k\geq 1} k(1-p)^k p \overset{(*)}{=} \frac{1}{p^2} \times (1-p)p = \frac{1-p}{p}$$

Auxilliary result $(*)$: for $|\alpha| < 1$,

$$\sum_{k\geq 1} k\alpha^{k-1} = \sum_{k\geq 1} \frac{d(\alpha^k)}{d\alpha} = \frac{d}{d\alpha} \sum_{k\geq 1} \alpha^k = \frac{d}{d\alpha}\left(\frac{\alpha}{1-\alpha}\right) = \frac{1}{(1-\alpha)^2}$$

$\square$

Intuitive explanation: “first-step analysis”

- Conditioned on the result of the first trial, we have

$$E(L) = p \times 0 + (1-p)E(1+L')$$

  where L' is the number of trials in addition to the first failed trial till reaching success.

- Note that L' and L have the same distribution, hence same mean. So

$$E(L) = (1-p)(1+E(L))$$

  Solving this equation gives $E(L) = (1-p)/p$.

Memoryless property of geometric distribution

$$P(L = k + \ell | L \geq k) = P(L = \ell), \quad k, \ell \geq 0$$

Interpretation: Having failed k times already, the probability that one fails another ℓ times is the same as failing ℓ times from the fresh start, as if the past is "forgotten."

[1] Alternatively, $P(L \geq k) = P(\text{first } k \text{ trials all failed}) = (1-p)^k$.

Memoryless property of geometric distribution

$$P(L = k + \ell | L \geq k) = P(L = \ell), \quad k, \ell \geq 0$$

Interpretation: Having failed k times already, the probability that one fails another ℓ times is the same as failing ℓ times from the fresh start, as if the past is "forgotten."

Proof.

$P(L \geq k) = \sum_{j \geq k} P(L = j) = \sum_{j \geq k} p(1-p)^j = (1-p)^k$.[1] Then

$$P(L = k + \ell | L \geq k) = \frac{P(L = k + \ell)}{P(L \geq k)} = \frac{p(1-p)^{k+\ell}}{(1-p)^k} = \underbrace{p(1-p)^\ell}_{P(L=\ell)}$$

$\square$

[1] Alternatively, $P(L \geq k) = P(\text{first } k \text{ trials all failed}) = (1-p)^k$.

Example: coupon collector

- There are n different coupons
- Each box of cereal contains one of n coupons chosen uniformly at random and independently
- A coupon aficionado keeps buying until all n coupons have been collected
- How many need to buy on average to complete the collection?

Example: coupon collector

- There are n different coupons
- Each box of cereal contains one of n coupons chosen uniformly at random and independently
- A coupon aficionado keeps buying until all n coupons have been collected
- How many need to buy on average to complete the collection?

Observations

- If extremely lucky, first n boxes contain all distinct coupons (Exercise: what is the chance?)

Example: coupon collector

- There are n different coupons
- Each box of cereal contains one of n coupons chosen uniformly at random and independently
- A coupon aficionado keeps buying until all n coupons have been collected
- How many need to buy on average to complete the collection?

Observations

- If extremely lucky, first n boxes contain all distinct coupons (Exercise: what is the chance?)
- Typically need to buy more than n because of repetitions. Question: how much more? $2n$? $10n$?

Example: coupon collector

- There are n different coupons
- Each box of cereal contains one of n coupons chosen uniformly at random and independently
- A coupon aficionado keeps buying until all n coupons have been collected
- How many need to buy on average to complete the collection?

Observations

- If extremely lucky, first n boxes contain all distinct coupons (Exercise: what is the chance?)
- Typically need to buy more than n because of repetitions. Question: how much more? $2n$? $10n$?

Let $X =$ number of boxes bought till completing the collection. Find $E(X)$.

Example: coupon collector

- Having collected $i-1$ distinct coupons, let X_i the number of boxes to buy till the next new one appears, $i = 1, 2, \ldots, n$.

Example: coupon collector

- Having collected $i-1$ distinct coupons, let X_i the number of boxes to buy till the next new one appears, $i = 1, 2, \ldots, n$.
- Then
$$X_i - 1 \sim \text{Geom}\underbrace{\left(1 - \frac{i-1}{n}\right)}_{P(\text{new})}, \quad E(X_i) = \frac{n}{n-i+1}$$

Example: coupon collector

- Having collected $i-1$ distinct coupons, let X_i the number of boxes to buy till the next new one appears, $i = 1, 2, \dots, n$.
- Then

$$X_i - 1 \sim \text{Geom}\underbrace{\left(1 - \frac{i-1}{n}\right)}_{P(\text{new})}, \quad E(X_i) = \frac{n}{n-i+1}$$

- Then $X = X_1 + \cdots + X_n$ and

$$\begin{aligned} E(X) &= E(X_1) + \cdots + E(X_{n-1}) \\ &= \underbrace{1 + \frac{n}{n-1} + \frac{n}{n-2} + \cdots + \frac{n}{1}}_{\text{harder and harder to find new coupon}} = n \sum_{i=1}^{n} \frac{1}{i} \overset{\text{[B-H A.8.4]}}{\approx} n \ln n \end{aligned}$$

On average need to buy an unbounded factor of $\ln n$ more (e.g. $n = 100, E(X) \approx 519$)

Example: coupon collector

- Having collected $i-1$ distinct coupons, let X_i the number of boxes to buy till the next new one appears, $i = 1, 2, \ldots, n$.
- Then
$$X_i - 1 \sim \text{Geom}\underbrace{\left(1 - \frac{i-1}{n}\right)}_{P(\text{new})}, \quad E(X_i) = \frac{n}{n-i+1}$$
- Then $X = X_1 + \cdots + X_n$ and
$$\begin{aligned} E(X) &= E(X_1) + \cdots + E(X_{n-1}) \\ &= \underbrace{1 + \frac{n}{n-1} + \frac{n}{n-2} + \cdots + \frac{n}{1}}_{\text{harder and harder to find new coupon}} = n \sum_{i=1}^{n} \frac{1}{i} \overset{\text{[B-H A.8.4]}}{\approx} n \ln n \end{aligned}$$

  On average need to buy an unbounded factor of $\ln n$ more (e.g. $n = 100, E(X) \approx 519$)
- In fact, $X_1, \cdots, X_n$ are independent (not needed for finding $E(X)$)

Summary

- Bern(p): 1 parameter. Indicator for the success of a single trial
- Bin(n, p): 2 parameters. Number of successes in n independent trials

$$\text{Bin}(n, p) = \text{sum of } n \text{ iid Bern}(p) \text{ random variables}$$

- Geom(p): 1 parameter. Number of failures till first success.

Binomial identities

$$\sum_{k=0}^{n} \binom{n}{k} x^k y^{n-k} = (x+y)^n$$

$$\sum_{i=0}^{k} \binom{n}{i} \binom{m}{k-i} = \binom{m+n}{k}$$

$$\binom{n}{n-k} = \binom{n}{k}$$

Exercise: c.f. B-H Sec 1.5 Story proofs

$$\binom{n-1}{k-1} + \binom{n-1}{k} = \binom{n}{k}$$

$$\sum_{k=0}^{n} \binom{n}{k}^2 = \binom{2n}{n}$$