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

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

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

\[ X_i = \begin{cases} 
1 & \text{ith trial succeeds} \\
0 & \text{ith trial fails} 
\end{cases} \]

• \(X_1, \ldots, X_n\) 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?
Let $X \sim \text{Bin}(n, p)$. How to find its PMF?

- $X$ takes values in $\{0, 1, \ldots, n\}$
Binomial PMF

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

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

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

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

- Recall binomial expansion:

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

- Substituting \(x = p\) and \(y = 1 - p\):

\[\sum_{k=0}^{n} \binom{n}{k} p^k (1 - p)^{n-k} = (p + (1 - p))^n = 1\]
Binomial PMF: $\text{Bin}(15, 1/2)$ vs $\text{Bin}(100, 1/2)$

\[ X \sim \text{Bin}(n, 1/2), \quad P(X = k) = \binom{n}{k} \frac{1}{2^n} \]
Binomial PMF

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

Bin(100, 1/10)
Binomial PMF

Bin\((100, 2/10)\)
Binomial PMF

Bin\(\left(100, \frac{3}{10}\right)\)
Binomial PMF

Bin\left(100, \frac{4}{10}\right)
Binomial PMF

Bin($100, 6/10$)
Binomial PMF

Bin(100, 7/10)
Binomial PMF

Bin$(100, 8/10)$
Binomial PMF

Bin(100, 9/10)
Binomial PMF

\[ \text{Bin}(100, 10/10) \]
Mean of Bin$(n, p)$

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

$E(X) = np$
Mean of Bin$(n, p)$

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

$E(X) = np$

Proof 1  Linearity of expectation

$$E(X) = E(X_1 + \ldots + X_n) = E(X_1) + \ldots + E(X_n) = np$$

Mode (most likely value) of Bin$(n, p)$ is $\lfloor (n+1)p \rfloor$: (Exercise)
**Mean of Bin$(n, p)$**

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

$E(X) = np$

**Proof 1** Linearity of expectation

\[
E(X) = E(X_1 + \ldots + X_n) = E\left( X_1 \right) + \ldots + E\left( X_n \right) = np
\]

**Proof 2** Direct calculation via PMF

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

where we applied $n\binom{n}{k} = k\binom{n-1}{k-1}$ [B-H Example 1.5.2]
Mean of $\text{Bin}(n, p)$

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

$E(X) = np$

Proof 1  Linearity of expectation

$$E(X) = E(X_1 + \ldots + X_n) = E\left( X_1 \right) + \ldots + E\left( X_n \right) = np$$

Proof 2  Direct calculation via PMF

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

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

Mode (most likely value) of $\text{Bin}(n, p)$ is $\lfloor (n + 1)p \rfloor$: (Exercise)
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.
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
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 $X + Y$?
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 $X \sim \text{Bin}(n, p)$. Let $Z = n - X$, which is a function of $X$. Then
Direct verification

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

$$P(Z = k) = P(X = n - k)$$

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

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

[B-H Example 1.5.1]

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

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

\[
P(S = k) = \sum_{i=0}^{k} P(X = i) P(Y = k-i) = \sum_{i=0}^{k} \binom{n}{i} p^i (1-p)^{n-i} \binom{m}{k-i} p^{k-i} (1-p)^{m-k+i}.
\]

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

\[\{z\}_{i=0}^{k} = \left(\frac{m+n}{k}\right)^{k}_{\binom{m+n}{k}}.\]
Direct verification

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

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

\[
= \sum_{i=0}^{k} \binom{n}{i} p^i (1 - p)^{n-i} \binom{m}{k-i} p^{k-i} (1 - p)^{m-k+i}
\]

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

[B-H Example 1.5.3]
Direct verification

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

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

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

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

$$= \binom{m+n}{k} \text{ [B-H Example 1.5.3]}$$

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 $\frac{1}{2}$ 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 $\frac{1}{2}$ 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

\[ P(Y = 4) = P(3 : 1) + P(1 : 3) \]
\[ = 2P(3 : 1) \]
\[ = 2P(\text{A wins 2 out of the first 3 games}) \times P(\text{A wins 4th game}) \]
\[ = 2 \left( \binom{3}{2} \frac{1}{2^3} \right) \times \frac{1}{2} \]
\[ = \frac{3}{8} \]

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

\[ P(Y = 5) = P(3 : 2) + P(2 : 3) \]
\[ = 2P(3 : 2) \]
\[ = 2P(A \text{ wins 2 out of the first 4 games}) \times P(A \text{ wins 5th game}) \]
\[ = 2 \left( \binom{4}{2} \right) \frac{1}{2^4} \times \frac{1}{2} \]
\[ = \frac{3}{8} \]

since the number of games won by A in the first 4 games \( \sim \text{Bin}(4, 1/2) \)
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 $L \sim \text{Geom}(p)$
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 $k \geq 0$,

$$P(L = k) = P(\text{fail first } k \text{ attempts, succeed in the } (k + 1)^{\text{th}})$$

$$= (1-p)^k p$$

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 $k \geq 0$,

\[
P(L = k) = P(\text{fail first } k \text{ attempts, succeed in the } (k+1)^{th})
\]

\[
= (1 - p)^k p
\]

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):

\[
\sum_{k \geq k_0} \alpha^k = \frac{\alpha^{k_0}}{1 - \alpha} = \frac{\text{first term}}{1 - \text{ratio}}
\]
Geom(0.5) vs Geom(0.2)
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
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.

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

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

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

$$\sum_{k \geq 1} k\alpha^{k-1} = \sum_{k \geq 1} \frac{d(\alpha^k)}{d\alpha} = \frac{d}{d\alpha} \sum_{k \geq 1} \alpha^k = \frac{d}{d\alpha} \left( \frac{\alpha}{1 - \alpha} \right) = \frac{1}{(1 - \alpha)^2}$$
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 \( \ell \) times is the same as failing \( \ell \) times from the fresh start, as if the past is “forgotten.”

\(^1\text{Alternatively, } P(L \geq k) = P(\text{first } k \text{ trials all failed}) = (1 - p)^k.\)
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 \( \ell \) times is the same as failing \( \ell \) times from the fresh start, as if the past is “forgotten.”

Proof.

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

Then

\[
P(L = k + \ell | L \geq k) = \frac{P(L = k + \ell)}{P(L \geq k)} = \frac{p(1 - p)^{k+\ell}}{(1 - p)^k} = \frac{p(1 - p)^\ell}{P(L = \ell)}
\]

\[\]

\[\] Alternately, \( 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

\[
X_i - 1 \sim \text{Geom} \left(1 - \frac{i - 1}{n}\right), \quad E(X_i) = \frac{n}{n - i + 1}
\]

\( P(\text{new}) \)

\[
\approx n \ln n
\]

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

- In fact, \( X_1, \ldots, X_n \) are independent (not needed for finding \( E(X_i) \)).
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

\[
X_i - 1 \sim \text{Geom} \left( 1 - \frac{i - 1}{n} \right), \quad E(X_i) = \frac{n}{n - i + 1}
\]

\( P(\text{new}) \)

• Then \( X = X_1 + \cdots + X_n \) and

\[
E(X) = E(X_1) + \cdots + E(X_{n-1})
\]

\[
= 1 + \frac{n}{n-1} + \frac{n}{n-2} + \cdots + \frac{n}{1} = n \sum_{i=1}^{n} \frac{1}{i} \] [B-H A.8.4] \approx n \ln n

harder and harder to find new coupon

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
  \[
  X_i - 1 \sim \text{Geom} \left( 1 - \frac{i - 1}{n} \right), \quad E(X_i) = \frac{n}{n - i + 1}
  \]

- Then \( X = X_1 + \cdots + X_n \) and
  \[
  E(X) = E(X_1) + \cdots + E(X_{n-1}) = 1 + \frac{n}{n-1} + \frac{n}{n-2} + \cdots + \frac{n}{1} = n \sum_{i=1}^{n} \frac{1}{i} \quad \approx \quad n \ln n
  \]
  On average need to buy an unbounded factor of \( \ln n \) more
  (e.g. \( n = 100, E(X) \approx 519 \))

- In fact, \( X_1, \ldots, X_n \) 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

\[ \binom{n-1}{k-1} + \binom{n-1}{k} = \binom{n}{k} \]

\[ \sum_{k=0}^{n} \binom{n}{k}^2 = \binom{2n}{n} \]