18.440: Lecture 11
Binomial random variables and repeated trials

Scott Sheffield

MIT
Outline

Bernoulli random variables

Properties: expectation and variance

More problems
Outline

Bernoulli random variables

Properties: expectation and variance

More problems
Bernoulli random variables

- Toss fair coin $n$ times. (Tosses are independent.) What is the probability of $k$ heads?
- Answer: $\binom{n}{k}/2^n$.
- What if coin has $p$ probability to be heads?
- Answer: $\binom{n}{k}p^k(1-p)^{n-k}$.
- Writing $q = 1 - p$, we can write this as $\binom{n}{k}p^kq^{n-k}$.
- Can use binomial theorem to show probabilities sum to one:
  - $1 = 1^n = (p + q)^n = \sum_{k=0}^{n} \binom{n}{k}p^kq^{n-k}$.
- Number of heads is **binomial random variable with parameters** $(n, p)$.
Examples

- Toss 6 fair coins. Let $X$ be number of heads you see. Then $X$ is binomial with parameters $(n, p)$ given by $(6, 1/2)$.

- Probability mass function for $X$ can be computed using the 6th row of Pascal’s triangle.

- If coin is biased (comes up heads with probability $p \neq 1/2$), we can still use the 6th row of Pascal’s triangle, but the probability that $X = i$ gets multiplied by $p^i(1 - p)^{n-i}$.
Other examples

- Room contains $n$ people. What is the probability that exactly $i$ of them were born on a Tuesday?
- **Answer:** use binomial formula $\binom{n}{i} p^i q^{n-i}$ with $p = 1/7$ and $q = 1 - p = 6/7$.
- Let $n = 100$. Compute the probability that nobody was born on a Tuesday.
- What is the probability that exactly 15 people were born on a Tuesday?
Outline

Bernoulli random variables

Properties: expectation and variance

More problems
Expectation

- Let $X$ be a binomial random variable with parameters $(n, p)$.
- What is $E[X]$?
- Direct approach: by definition of expectation,
  $$E[X] = \sum_{i=0}^{n} P\{X = i\}i.$$ 
- What happens if we modify the $n$th row of Pascal’s triangle by multiplying the $i$ term by $i$?
- For example, replace the 5th row $(1, 5, 10, 10, 5, 1)$ by $(0, 5, 20, 30, 20, 5)$. Does this remind us of an earlier row in the triangle?
- Perhaps the prior row $(1, 4, 6, 4, 1)$?
Useful Pascal’s triangle identity

Recall that \( \binom{n}{i} = \frac{n \times (n-1) \times \ldots \times (n-i+1)}{i \times (i-1) \times \ldots \times 1} \). This implies a simple but important identity: \( i \binom{n}{i} = n \binom{n-1}{i-1} \).

Using this identity (and \( q = 1 - p \)), we can write

\[
E[X] = \sum_{i=0}^{n} i \binom{n}{i} p^i q^{n-i} = \sum_{i=1}^{n} n \binom{n-1}{i-1} p^i q^{n-i}.
\]

Rewrite this as \( E[X] = np \sum_{i=1}^{n} \binom{n-1}{i-1} p^{(i-1)} q^{(n-1)-(i-1)} \).

Substitute \( j = i - 1 \) to get

\[
E[X] = np \sum_{j=0}^{n-1} \binom{n-1}{j} p^j q^{(n-1)-j} = np(p + q)^{n-1} = np.
\]
Let $X$ be a binomial random variable with parameters $(n, p)$. Here is another way to compute $E[X]$.

Think of $X$ as representing number of heads in $n$ tosses of coin that is heads with probability $p$.

Write $X = \sum_{j=1}^{n} X_j$, where $X_j$ is 1 if the $j$th coin is heads, 0 otherwise.

In other words, $X_j$ is the number of heads (zero or one) on the $j$th toss.

Note that $E[X_j] = p \cdot 1 + (1 - p) \cdot 0 = p$ for each $j$.

Conclude by additivity of expectation that

$$E[X] = \sum_{j=1}^{n} E[X_j] = \sum_{j=1}^{n} p = np.$$
Let $X$ be binomial $(n, p)$ and fix $k \geq 1$. What is $E[X^k]$?

Recall identity: $i \binom{n}{i} = n \binom{n-1}{i-1}$.

Generally, $E[X^k]$ can be written as

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

Identity gives

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

$$np \sum_{j=0}^{n-1} \binom{n-1}{j} p^{j} (1 - p)^{n-1-j} (j + 1)^{k-1}.$$ 

Thus $E[X^k] = E[(Y + 1)^{k-1}]$ where $Y$ is binomial with parameters $(n - 1, p)$. 

Interesting moment computation
Computing the variance

- Let \( X \) be binomial \((n, p)\). What is \( E[X] \)?
- We know \( E[X] = np \).
- We computed identity \( E[X^k] = E[(Y + 1)^{k-1}] \) where \( Y \) is binomial with parameters \((n - 1, p)\).
- In particular \( E[X^2] = npE[Y + 1] = np[(n - 1)p + 1] \).
- So \( \text{Var}[X] = E[X^2] - E[X]^2 = np(n - 1)p + np - (np)^2 = np(1 - p) = npq \), where \( q = 1 - p \).
- Commit to memory: variance of binomial \((n, p)\) random variable is \( npq \).
- This is \( n \) times the variance you’d get with a single coin. Coincidence?
Compute variance with decomposition trick

- $X = \sum_{j=1}^{n} X_j$, so
  $E[X^2] = E[\sum_{i=1}^{n} X_i \sum_{j=1}^{n} X_j] = \sum_{i=1}^{n} \sum_{j=1}^{n} E[X_i X_j]$

- $E[X_i X_j]$ is $p$ if $i = j$, $p^2$ otherwise.

- $\sum_{i=1}^{n} \sum_{j=1}^{n} E[X_i X_j]$ has $n$ terms equal to $p$ and $(n - 1)n$ terms equal to $p^2$.

- So $E[X^2] = np + (n - 1)np^2 = np + (np)^2 - np^2$.

- Thus
Outline

Bernoulli random variables

Properties: expectation and variance

More problems
More examples

- An airplane seats 200, but the airline has sold 205 tickets. Each person, independently, has a .05 chance of not showing up for the flight. What is the probability that more than 200 people will show up for the flight?

- In a 100 person senate, forty people always vote for the Republicans’ position, forty people always for the Democrats’ position and 20 people just toss a coin to decide which way to vote. What is the probability that a given vote is tied?

- You invite 50 friends to a party. Each one, independently, has a 1/3 chance of showing up. That is the probability that more than 25 people will show up?