18.600: Lecture 14

More discrete random variables

Scott Sheffield

MIT

Outline

Geometric random variables

Negative binomial random variables

Problems

Outline

Geometric random variables

Negative binomial random variables

Problems

Geometric random variables

- Consider an infinite sequence of independent tosses of a coin that comes up heads with probability p.

Geometric random variables

- Consider an infinite sequence of independent tosses of a coin that comes up heads with probability p.
- Let X be such that the first heads is on the Xth toss.

Geometric random variables

- Consider an infinite sequence of independent tosses of a coin that comes up heads with probability p.
- Let X be such that the first heads is on the Xth toss.
- For example, if the coin sequence is $T, T, H, T, H, T, \ldots$ then $X = 3$.

Geometric random variables

- Consider an infinite sequence of independent tosses of a coin that comes up heads with probability p.
- Let X be such that the first heads is on the Xth toss.
- For example, if the coin sequence is $T, T, H, T, H, T, \ldots$ then $X = 3$.
- Then X is a random variable. What is $P\{X = k\}$?

Geometric random variables

- Consider an infinite sequence of independent tosses of a coin that comes up heads with probability p.
- Let X be such that the first heads is on the Xth toss.
- For example, if the coin sequence is $T, T, H, T, H, T, \ldots$ then $X = 3$.
- Then X is a random variable. What is $P\{X = k\}$?
- Answer: $P\{X = k\} = (1-p)^{k-1}p = q^{k-1}p$, where $q = 1 - p$ is tails probability.

Geometric random variables

- Consider an infinite sequence of independent tosses of a coin that comes up heads with probability p.
- Let X be such that the first heads is on the Xth toss.
- For example, if the coin sequence is $T, T, H, T, H, T, \ldots$ then $X = 3$.
- Then X is a random variable. What is $P\{X = k\}$?
- Answer: $P\{X = k\} = (1-p)^{k-1}p = q^{k-1}p$, where $q = 1 - p$ is tails probability.
- Can you prove directly that these probabilities sum to one?

Geometric random variables

- Consider an infinite sequence of independent tosses of a coin that comes up heads with probability p.
- Let X be such that the first heads is on the Xth toss.
- For example, if the coin sequence is $T, T, H, T, H, T, \ldots$ then $X = 3$.
- Then X is a random variable. What is $P\{X = k\}$?
- Answer: $P\{X = k\} = (1-p)^{k-1}p = q^{k-1}p$, where $q = 1 - p$ is tails probability.
- Can you prove directly that these probabilities sum to one?
- Say X is a **geometric random variable with parameter** p.

Geometric random variable expectation

- Let X be a geometric with parameter p, i.e., $P\{X = k\} = (1-p)^{k-1}p = q^{k-1}p$ for $k \geq 1$.

Geometric random variable expectation

- Let X be a geometric with parameter p, i.e., $P\{X = k\} = (1-p)^{k-1}p = q^{k-1}p$ for $k \geq 1$.
- What is $E[X]$?

Geometric random variable expectation

- Let X be a geometric with parameter p, i.e., $P\{X = k\} = (1-p)^{k-1}p = q^{k-1}p$ for $k \geq 1$.
- What is $E[X]$?
- By definition $E[X] = \sum_{k=1}^{\infty} q^{k-1}pk$.

Geometric random variable expectation

- Let X be a geometric with parameter p, i.e., $P\{X = k\} = (1-p)^{k-1}p = q^{k-1}p$ for $k \geq 1$.
- What is $E[X]$?
- By definition $E[X] = \sum_{k=1}^{\infty} q^{k-1}pk$.
- There's a trick to computing sums like this.

Geometric random variable expectation

- Let X be a geometric with parameter p, i.e., $P\{X = k\} = (1-p)^{k-1}p = q^{k-1}p$ for $k \geq 1$.
- What is $E[X]$?
- By definition $E[X] = \sum_{k=1}^{\infty} q^{k-1}pk$.
- There's a trick to computing sums like this.
- Note $E[X-1] = \sum_{k=1}^{\infty} q^{k-1}p(k-1)$. Setting $j = k-1$, we have $E[X-1] = q\sum_{j=0}^{\infty} q^{j-1}pj = qE[X]$.

Geometric random variable expectation

- Let X be a geometric with parameter p, i.e., $P\{X = k\} = (1-p)^{k-1}p = q^{k-1}p$ for $k \geq 1$.
- What is $E[X]$?
- By definition $E[X] = \sum_{k=1}^{\infty} q^{k-1}pk$.
- There's a trick to computing sums like this.
- Note $E[X-1] = \sum_{k=1}^{\infty} q^{k-1}p(k-1)$. Setting $j = k-1$, we have $E[X-1] = q\sum_{j=0}^{\infty} q^{j-1}pj = qE[X]$.
- Kind of makes sense. $X-1$ is "number of extra tosses after first." Given first coin heads (probability p), $X-1$ is 0. Given first coin tails (probability q), conditional law of $X-1$ is geometric with parameter p. In latter case, conditional expectation of $X-1$ is same as a priori expectation of X.

Geometric random variable expectation

- Let X be a geometric with parameter p, i.e., $P\{X = k\} = (1-p)^{k-1}p = q^{k-1}p$ for $k \geq 1$.
- What is $E[X]$?
- By definition $E[X] = \sum_{k=1}^{\infty} q^{k-1}pk$.
- There's a trick to computing sums like this.
- Note $E[X-1] = \sum_{k=1}^{\infty} q^{k-1}p(k-1)$. Setting $j = k-1$, we have $E[X-1] = q\sum_{j=0}^{\infty} q^{j-1}pj = qE[X]$.
- Kind of makes sense. $X - 1$ is "number of extra tosses after first." Given first coin heads (probability p), $X - 1$ is 0. Given first coin tails (probability q), conditional law of $X - 1$ is geometric with parameter p. In latter case, conditional expectation of $X - 1$ is same as a priori expectation of X.
- Thus $E[X] - 1 = E[X-1] = p \cdot 0 + qE[X] = qE[X]$ and solving for $E[X]$ gives $E[X] = 1/(1-q) = 1/p$.

Geometric random variable variance

- Let X be a geometric random variable with parameter p. Then $P\{X = k\} = q^{k-1}p$.

Geometric random variable variance

- Let X be a geometric random variable with parameter p. Then $P\{X = k\} = q^{k-1}p$.
- What is $E[X^2]$?

Geometric random variable variance

- Let X be a geometric random variable with parameter p. Then $P\{X = k\} = q^{k-1}p$.
- What is $E[X^2]$?
- By definition $E[X^2] = \sum_{k=1}^{\infty} q^{k-1}pk^2$.

Geometric random variable variance

- Let X be a geometric random variable with parameter p. Then $P\{X = k\} = q^{k-1}p$.
- What is $E[X^2]$?
- By definition $E[X^2] = \sum_{k=1}^{\infty} q^{k-1}pk^2$.
- Let's try to come up with a similar trick.

Geometric random variable variance

- Let X be a geometric random variable with parameter p. Then $P\{X = k\} = q^{k-1}p$.
- What is $E[X^2]$?
- By definition $E[X^2] = \sum_{k=1}^{\infty} q^{k-1}pk^2$.
- Let's try to come up with a similar trick.
- Note $E[(X-1)^2] = \sum_{k=1}^{\infty} q^{k-1}p(k-1)^2$. Setting $j = k-1$, we have $E[(X-1)^2] = q\sum_{j=0}^{\infty} q^{j-1}pj^2 = qE[X^2]$.

Geometric random variable variance

- Let X be a geometric random variable with parameter p. Then $P\{X = k\} = q^{k-1}p$.
- What is $E[X^2]$?
- By definition $E[X^2] = \sum_{k=1}^{\infty} q^{k-1}pk^2$.
- Let's try to come up with a similar trick.
- Note $E[(X-1)^2] = \sum_{k=1}^{\infty} q^{k-1}p(k-1)^2$. Setting $j = k-1$, we have $E[(X-1)^2] = q\sum_{j=0}^{\infty} q^{j-1}pj^2 = qE[X^2]$.
- Thus $E[(X-1)^2] = E[X^2 - 2X + 1] = E[X^2] - 2E[X] + 1 = E[X^2] - 2/p + 1 = qE[X^2]$.

Geometric random variable variance

- Let X be a geometric random variable with parameter p. Then $P\{X = k\} = q^{k-1}p$.
- What is $E[X^2]$?
- By definition $E[X^2] = \sum_{k=1}^{\infty} q^{k-1}pk^2$.
- Let's try to come up with a similar trick.
- Note $E[(X-1)^2] = \sum_{k=1}^{\infty} q^{k-1}p(k-1)^2$. Setting $j = k-1$, we have $E[(X-1)^2] = q\sum_{j=0}^{\infty} q^{j-1}pj^2 = qE[X^2]$.
- Thus $E[(X-1)^2] = E[X^2 - 2X + 1] = E[X^2] - 2E[X] + 1 = E[X^2] - 2/p + 1 = qE[X^2]$.
- Solving for $E[X^2]$ gives $(1-q)E[X^2] = pE[X^2] = 2/p - 1$, so $E[X^2] = (2-p)/p^2$.

Geometric random variable variance

- Let X be a geometric random variable with parameter p. Then $P\{X = k\} = q^{k-1}p$.
- What is $E[X^2]$?
- By definition $E[X^2] = \sum_{k=1}^{\infty} q^{k-1}pk^2$.
- Let's try to come up with a similar trick.
- Note $E[(X-1)^2] = \sum_{k=1}^{\infty} q^{k-1}p(k-1)^2$. Setting $j = k-1$, we have $E[(X-1)^2] = q\sum_{j=0}^{\infty} q^{j-1}pj^2 = qE[X^2]$.
- Thus $E[(X-1)^2] = E[X^2 - 2X + 1] = E[X^2] - 2E[X] + 1 = E[X^2] - 2/p + 1 = qE[X^2]$.
- Solving for $E[X^2]$ gives $(1-q)E[X^2] = pE[X^2] = 2/p - 1$, so $E[X^2] = (2-p)/p^2$.
- $\mathrm{Var}[X] = (2-p)/p^2 - 1/p^2 = (1-p)/p^2 = 1/p^2 - 1/p = q/p^2$.

Example

- Toss die repeatedly. Say we get 6 for first time on Xth toss.

Example

- Toss die repeatedly. Say we get 6 for first time on Xth toss.
- What is $P\{X = k\}$?

Example

- Toss die repeatedly. Say we get 6 for first time on Xth toss.
- What is $P\{X = k\}$?
- Answer: $(5/6)^{k-1}(1/6)$.

Example

- Toss die repeatedly. Say we get 6 for first time on Xth toss.
- What is $P\{X = k\}$?
- Answer: $(5/6)^{k-1}(1/6)$.
- What is $E[X]$?

Example

- Toss die repeatedly. Say we get 6 for first time on Xth toss.
- What is $P\{X = k\}$?
- Answer: $(5/6)^{k-1}(1/6)$.
- What is $E[X]$?
- Answer: 6.

Example

- Toss die repeatedly. Say we get 6 for first time on Xth toss.
- What is $P\{X = k\}$?
- Answer: $(5/6)^{k-1}(1/6)$.
- What is $E[X]$?
- Answer: 6.
- What is $\mathrm{Var}[X]$?

Example

- Toss die repeatedly. Say we get 6 for first time on Xth toss.
- What is $P\{X = k\}$?
- Answer: $(5/6)^{k-1}(1/6)$.
- What is $E[X]$?
- Answer: 6.
- What is $\mathrm{Var}[X]$?
- Answer: $1/p^2 - 1/p = 36 - 6 = 30$.

Example

- Toss die repeatedly. Say we get 6 for first time on Xth toss.
- What is $P\{X = k\}$?
- Answer: $(5/6)^{k-1}(1/6)$.
- What is $E[X]$?
- Answer: 6.
- What is $\mathrm{Var}[X]$?
- Answer: $1/p^2 - 1/p = 36 - 6 = 30$.
- Takes $1/p$ coin tosses on average to see a heads.

Outline

Outline

Negative binomial random variables

- Consider an infinite sequence of independent tosses of a coin that comes up heads with probability p.

Negative binomial random variables

- Consider an infinite sequence of independent tosses of a coin that comes up heads with probability p.
- Let X be such that the rth heads is on the Xth toss.

Negative binomial random variables

- Consider an infinite sequence of independent tosses of a coin that comes up heads with probability p.
- Let X be such that the rth heads is on the Xth toss.
- For example, if $r = 3$ and the coin sequence is $T, T, H, H, T, T, H, T, T, \ldots$ then $X = 7$.

Negative binomial random variables

- Consider an infinite sequence of independent tosses of a coin that comes up heads with probability p.
- Let X be such that the rth heads is on the Xth toss.
- For example, if $r = 3$ and the coin sequence is $T, T, H, H, T, T, H, T, T, \ldots$ then $X = 7$.
- Then X is a random variable. What is $P\{X = k\}$?

Negative binomial random variables

- Consider an infinite sequence of independent tosses of a coin that comes up heads with probability p.
- Let X be such that the rth heads is on the Xth toss.
- For example, if $r = 3$ and the coin sequence is $T, T, H, H, T, T, H, T, T, \ldots$ then $X = 7$.
- Then X is a random variable. What is $P\{X = k\}$?
- Answer: need exactly $r - 1$ heads among first $k - 1$ tosses and a heads on the kth toss.

Negative binomial random variables

- Consider an infinite sequence of independent tosses of a coin that comes up heads with probability p.
- Let X be such that the rth heads is on the Xth toss.
- For example, if $r = 3$ and the coin sequence is $T, T, H, H, T, T, H, T, T, \ldots$ then $X = 7$.
- Then X is a random variable. What is $P\{X = k\}$?
- Answer: need exactly $r - 1$ heads among first $k - 1$ tosses and a heads on the kth toss.
- So $P\{X = k\} = \binom{k-1}{r-1} p^{r-1}(1-p)^{k-r} p$. Can you prove these sum to 1?

Negative binomial random variables

- Consider an infinite sequence of independent tosses of a coin that comes up heads with probability p.
- Let X be such that the rth heads is on the Xth toss.
- For example, if $r = 3$ and the coin sequence is $T, T, H, H, T, T, H, T, T, \ldots$ then $X = 7$.
- Then X is a random variable. What is $P\{X = k\}$?
- Answer: need exactly $r - 1$ heads among first $k - 1$ tosses and a heads on the kth toss.
- So $P\{X = k\} = \binom{k-1}{r-1} p^{r-1}(1-p)^{k-r} p$. Can you prove these sum to 1?
- Call X **negative binomial random variable with parameters** (r, p).

Expectation of binomial random variable

- Consider an infinite sequence of independent tosses of a coin that comes up heads with probability p.

Expectation of binomial random variable

- Consider an infinite sequence of independent tosses of a coin that comes up heads with probability p.
- Let X be such that the rth heads is on the Xth toss.

Expectation of binomial random variable

- Consider an infinite sequence of independent tosses of a coin that comes up heads with probability p.
- Let X be such that the rth heads is on the Xth toss.
- Then X is a **negative binomial random variable with parameters** (r, p).

Expectation of binomial random variable

- Consider an infinite sequence of independent tosses of a coin that comes up heads with probability p.
- Let X be such that the rth heads is on the Xth toss.
- Then X is a **negative binomial random variable with parameters** (r, p).
- What is $E[X]$?

Expectation of binomial random variable

- Consider an infinite sequence of independent tosses of a coin that comes up heads with probability p.
- Let X be such that the rth heads is on the Xth toss.
- Then X is a **negative binomial random variable with parameters** (r, p).
- What is $E[X]$?
- Write $X = X_1 + X_2 + \ldots + X_r$ where X_k is number of tosses (following $(k-1)$th head) required to get kth head. Each X_k is geometric with parameter p.

Expectation of binomial random variable

- Consider an infinite sequence of independent tosses of a coin that comes up heads with probability p.
- Let X be such that the rth heads is on the Xth toss.
- Then X is a **negative binomial random variable with parameters** (r, p).
- What is $E[X]$?
- Write $X = X_1 + X_2 + \ldots + X_r$ where X_k is number of tosses (following $(k-1)$th head) required to get kth head. Each X_k is geometric with parameter p.
- So $E[X] = E[X_1 + X_2 + \ldots + X_r] = E[X_1] + E[X_2] + \ldots + E[X_r] = r/p$.

Expectation of binomial random variable

- Consider an infinite sequence of independent tosses of a coin that comes up heads with probability p.
- Let X be such that the rth heads is on the Xth toss.
- Then X is a **negative binomial random variable with parameters** (r, p).
- What is $E[X]$?
- Write $X = X_1 + X_2 + \ldots + X_r$ where X_k is number of tosses (following $(k-1)$th head) required to get kth head. Each X_k is geometric with parameter p.
- So $E[X] = E[X_1 + X_2 + \ldots + X_r] = E[X_1] + E[X_2] + \ldots + E[X_r] = r/p$.
- How about $\mathrm{Var}[X]$?

Expectation of binomial random variable

- Consider an infinite sequence of independent tosses of a coin that comes up heads with probability p.
- Let X be such that the rth heads is on the Xth toss.
- Then X is a **negative binomial random variable with parameters** (r, p).
- What is $E[X]$?
- Write $X = X_1 + X_2 + \ldots + X_r$ where X_k is number of tosses (following $(k-1)$th head) required to get kth head. Each X_k is geometric with parameter p.
- So $E[X] = E[X_1 + X_2 + \ldots + X_r] = E[X_1] + E[X_2] + \ldots + E[X_r] = r/p$.
- How about $\mathrm{Var}[X]$?
- Turns out that $\mathrm{Var}[X] = \mathrm{Var}[X_1] + \mathrm{Var}[X_2] + \ldots + \mathrm{Var}[X_r]$. So $\mathrm{Var}[X] = rq/p^2$.

Outline

Outline

Problems

- Nate and Natasha have beautiful new baby. Each minute with .01 probability (independent of all else) baby cries.

Problems

- Nate and Natasha have beautiful new baby. Each minute with .01 probability (independent of all else) baby cries.
- **Additivity of expectation:** How many times do they expect the baby to cry between 9 p.m. and 6 a.m.?

Problems

- Nate and Natasha have beautiful new baby. Each minute with .01 probability (independent of all else) baby cries.
- **Additivity of expectation:** How many times do they expect the baby to cry between 9 p.m. and 6 a.m.?
- **Geometric random variables:** What's the probability baby is quiet from midnight to three, then cries at exactly three?

Problems

- Nate and Natasha have beautiful new baby. Each minute with .01 probability (independent of all else) baby cries.
- **Additivity of expectation:** How many times do they expect the baby to cry between 9 p.m. and 6 a.m.?
- **Geometric random variables:** What's the probability baby is quiet from midnight to three, then cries at exactly three?
- **Geometric random variables:** What's the probability baby is quiet from midnight to three?

Problems

- Nate and Natasha have beautiful new baby. Each minute with .01 probability (independent of all else) baby cries.
- **Additivity of expectation:** How many times do they expect the baby to cry between 9 p.m. and 6 a.m.?
- **Geometric random variables:** What's the probability baby is quiet from midnight to three, then cries at exactly three?
- **Geometric random variables:** What's the probability baby is quiet from midnight to three?
- **Negative binomial:** Probability fifth cry is at midnight?

Problems

- Nate and Natasha have beautiful new baby. Each minute with .01 probability (independent of all else) baby cries.
- **Additivity of expectation:** How many times do they expect the baby to cry between 9 p.m. and 6 a.m.?
- **Geometric random variables:** What's the probability baby is quiet from midnight to three, then cries at exactly three?
- **Geometric random variables:** What's the probability baby is quiet from midnight to three?
- **Negative binomial:** Probability fifth cry is at midnight?
- **Negative binomial expectation:** How many minutes do I expect to wait until the fifth cry?

Problems

- Nate and Natasha have beautiful new baby. Each minute with .01 probability (independent of all else) baby cries.
- **Additivity of expectation:** How many times do they expect the baby to cry between 9 p.m. and 6 a.m.?
- **Geometric random variables:** What's the probability baby is quiet from midnight to three, then cries at exactly three?
- **Geometric random variables:** What's the probability baby is quiet from midnight to three?
- **Negative binomial:** Probability fifth cry is at midnight?
- **Negative binomial expectation:** How many minutes do I expect to wait until the fifth cry?
- **Poisson approximation:** Approximate the probability there are exactly five cries during the night.

Problems

- Nate and Natasha have beautiful new baby. Each minute with .01 probability (independent of all else) baby cries.
- **Additivity of expectation:** How many times do they expect the baby to cry between 9 p.m. and 6 a.m.?
- **Geometric random variables:** What's the probability baby is quiet from midnight to three, then cries at exactly three?
- **Geometric random variables:** What's the probability baby is quiet from midnight to three?
- **Negative binomial:** Probability fifth cry is at midnight?
- **Negative binomial expectation:** How many minutes do I expect to wait until the fifth cry?
- **Poisson approximation:** Approximate the probability there are exactly five cries during the night.
- **Exponential random variable approximation:** Approximate probability baby quiet all night.

More fun problems

- Suspose two soccer teams play each other. One team's number of points is Poisson with parameter λ_1 and other's is independently Poisson with parameter λ_2. (You can google "soccer" and "Poisson" to see the academic literature on the use of Poisson random variables to model soccer scores.) Using Mathematica (or similar software) compute the probability that the first team wins if $\lambda_1 = 2$ and $\lambda_2 = 1$. What if $\lambda_1 = 2$ and $\lambda_2 = .5$?

More fun problems

- Suspose two soccer teams play each other. One team's number of points is Poisson with parameter λ_1 and other's is independently Poisson with parameter λ_2. (You can google "soccer" and "Poisson" to see the academic literature on the use of Poisson random variables to model soccer scores.) Using Mathematica (or similar software) compute the probability that the first team wins if $\lambda_1 = 2$ and $\lambda_2 = 1$. What if $\lambda_1 = 2$ and $\lambda_2 = .5$?
- Imagine you start with the number 60. Then you toss a fair coin to decide whether to add 5 to your number or subtract 5 from it. Repeat this process with independent coin tosses until the number reaches 100 or 0. What is the *expected* number of tosses needed until this occurs?