More Common Discrete RVs

Announcements

Please sign up for small group discussion Office Hours for more help:
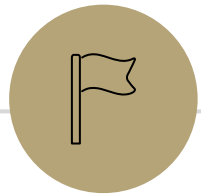
https://forms.gle/hkXnv1C2dHtBDFL58

I will make assignments sometime over the weekend.
Fill this out by Friday.

Real World 1 due tonight! Q 1.2a info - Ed Post

Think of problems / concepts that you want to review during section.

There is a section "midterm review" handout for preparation!

Negative Binomial Distribution

Scenario: Negative Binomial

You're playing a carnival game, and there are r little kids nearby who all want a stuffed animal. You can win a single game (and thus win one stuffed animal) with probability p (independently each time) How many times will you need to play the game before every kid gets their toy?

More generally, run independent trials with probability p. How many trials do you need for r successes?

Activity

More generally, run independent trials with probability p. How many trials do you need for r successes?

What's the pmf?

What's the expectation and variance? (hint: linearity)

Negative Binomial Analysis: PMF

What's the pmf? Well, how would we know $X = k$?

Of the first $k - 1$ trials, $r - 1$ must be successes.
And trial k must be a success.

That first part is a lot like a binomial!

It's the $p_Y(r - 1)$ where $Y \sim \text{Bin}(k - 1, p)$

First part gives $\binom{k-1}{r-1}(1 - p)^{k-1-(r-1)}p^{r-1} = \binom{k-1}{r-1}(1 - p)^{k-r}p^{r-1}$

Second part, multiply by p

Total: $p_X(k) = \binom{k-1}{r-1}(1 - p)^{k-r}p^r \quad for\ k \in \mathbb{Z}, r \leq k$

Negative Binomial Analysis: Expectation

What about the expectation?

To see r successes:

We flip until we see success 1.

Then flip until success 2.

… Flip until success r.

The total number of flips is…the sum of geometric random variables!

Negative Binomial Analysis: Expectation

Let $Z_1, Z_2, \ldots, Z_r$ be independent copies of $\mathrm{Geo}(p)$

Z_i are called "independent and identically distributed" or "i.i.d.'

Because they are independent...and have identical pmfs.

$X \sim \mathrm{NegBin}(r, p)$ $X = Z_1 + Z_2 + \cdots + Z_r$.

$$\mathbb{E}[X] = \mathbb{E}[Z_1 + Z_2 + \cdots Z_r] = \mathbb{E}[Z_1] + \mathbb{E}[Z_2] + \cdots + \mathbb{E}[Z_r] = r \cdot \frac{1}{p}$$

Negative Binomial Analysis: Variance

Let $Z_1, Z_2, \ldots, Z_r$ be independent copies of $\mathrm{Geo}(p)$

$X \sim \mathrm{NegBin}(r, p)$ $X = Z_1 + Z_2 + \cdots + Z_r$.

$\mathrm{Var}(X) = \mathrm{Var}(Z_1 + Z_2 + \cdots + Z_r)$

Up until now we've just used the observation that $X = Z_1 + \cdots + Z_r$.

$= \mathrm{Var}(Z_1) + \mathrm{Var}(Z_2) + \cdots + \mathrm{Var}(Z_r)$ because the Z_i are independent.

$= r \cdot \dfrac{1-p}{p^2}$

Negative Binomial

$X \sim \text{NegBin}(r, p)$

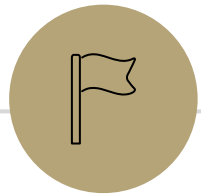Parameters: r: the number of successes needed, p the probability of success in a single trial

X is the number of trials needed to get the rth success.

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

$F_X(k)$ is ugly, don't bother with it.

$$\mathbb{E}[X] = \frac{r}{p}$$

$$\text{Var}(X) = \frac{r(1-p)}{p^2}$$

Hypergeometric Distribution

Scenario: Hypergeometric

You have an urn with N balls, of which K are purple. You are going to draw balls out of the urn **without** replacement.

If you draw out n balls, what is the probability you see k purple ones?

Hypergeometric: Analysis

You have an urn with N balls, of which K are purple. You are going to draw balls out of the urn **without** replacement.

If you draw out n balls, what is the probability you see k purple ones?

Of the K purple, we draw out $k.$ Choose which k will be drawn

Of the $N - K$ other balls, we will draw out $n - k$, choose which $(n - k)$ will be drawn.

Sample space all subsets of size n

$$\frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}$$

Hypergeometric: Analysis

$X = D_1 + D_2 + \cdots + D_n$

Where D_i is the indicator that draw i is purple.

D_1 is 1 with probability K/N.

What about D_2?

$$\mathbb{P}(D_2 = 1) = \frac{K-1}{N-1} \cdot \frac{K}{N} + \frac{K}{N-1} \cdot \frac{K-N}{N} = \frac{K(K-N+K-1)}{N(N-1)} = \frac{K}{N}$$

In general, $\mathbb{P}(D_i = 1) = \dfrac{K}{N}$

It might feel counterintuitive, but it's true!

Hypergeometric: Analysis

$\mathbb{E}[X]$

$= \mathbb{E}[D_1 + \cdots D_n] = \mathbb{E}[D_1] + \cdots + \mathbb{E}[D_n] = n \cdot \dfrac{K}{N}$

Can we do the same for variance?

No! The D_i are dependent. Even if they have the same probability.

Hypergeometric Distribution

$X \sim \mathrm{HypGeo}(N, K, n)$

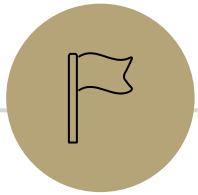Parameters: A total of N balls in an urn, of which K are successes. Draw n balls without replacement.

X is the number of success balls drawn.

$$p_X(k) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}$$

$$\mathbb{E}[X] = \frac{nK}{N}$$

$$\mathrm{Var}(X) = n \cdot \frac{K}{N} \cdot \frac{N-K}{N} \cdot \frac{N-n}{N-1}$$

Poisson Distribution

The Poisson Distribution

A new kind of random variable.

We use a Poisson distribution when:

We're trying to count the number of times something happens in some interval of time.

We know the average number that happen (i.e. the expectation)

Each occurrence is independent of the others.

There are a VERY large number of "potential sources" for those events, few of which happen.

The Poisson Distribution

Classic applications:

How many traffic accidents occur in Seattle in a day?

How many major earthquakes occur in a year (not including aftershocks)?

How many customers visit a bakery in an hour?

Why not just use counting coin flips?

What are the flips…the number of cars? Every person who might visit the bakery? There are way too many of these to count exactly or think about dependency between. But a Poisson might accurately model what's happening.

It's a model

By modeling choice, we mean that we're choosing math that we think represents the real world as best as possible

Is every traffic accident really independent?

Not *really,* one causes congestion, which causes angrier drivers. Or both might be caused by bad weather/more cars on the road.

But we assume they are (because the dependence is so weak that the model is useful).

Poisson Distribution

$X \sim \text{Poi}(\lambda)$

Let λ be the average number of incidents in a time interval.

X is the number of incidents seen in a particular interval.

Support $\mathbb{N}$

$$p_X(k) = \frac{\lambda^k e^{-\lambda}}{k!} \text{ (for } k \in \mathbb{N})$$
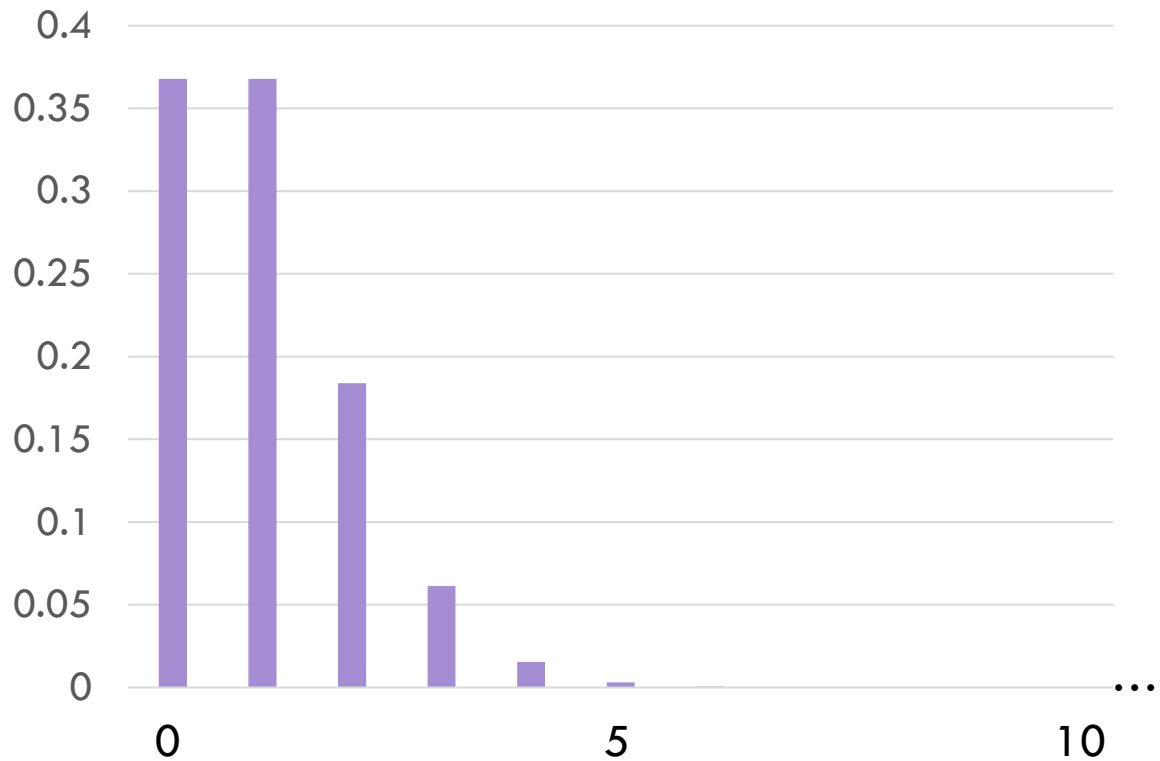
$$F_X(k) = e^{-\lambda} \sum_{i=0}^{\lfloor k \rfloor} \frac{\lambda^i}{i!}$$
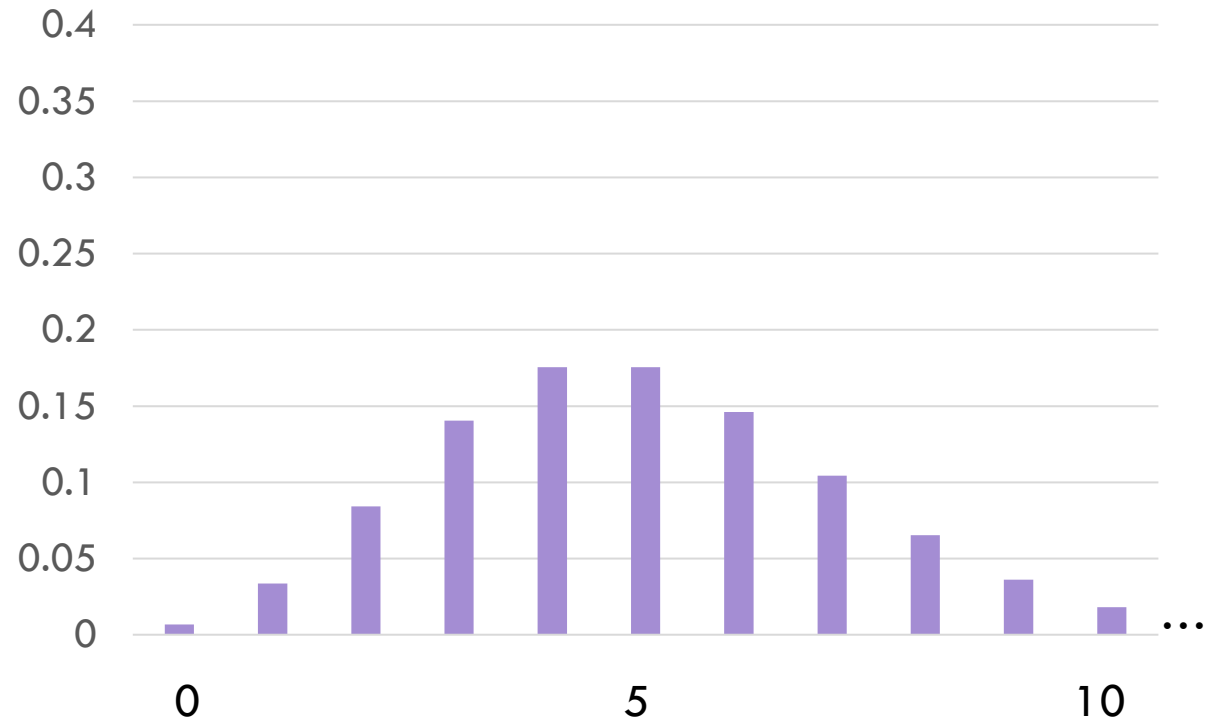
$$\mathbb{E}[X] = \lambda$$

$$\text{Var}(X) = \lambda$$

Some Sample PMFs

PMF for Poisson with lambda=1

PMF for Poisson with lambda=5

Let's take a closer look at that pmf

$$p_X(k) = \frac{\lambda^k e^{-\lambda}}{k!} \text{ (for } k \in \mathbb{N})$$

If this is a real PMF, it should sum to 1.
Let's check that to understand the PMF a little better.

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

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

$$\boxed{\begin{array}{c} Taylor\ Series\ for\ e^x \\ \displaystyle\sum_{k=0}^{\infty} \frac{x^k}{k!} = e^x \end{array}}$$

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

Let's check something...the expectation

$$\mathbb{E}[X] = \sum_{k=0}^{\infty} k \cdot e^{-\lambda} \frac{\lambda^k}{k!}$$

$$= \sum_{k=1}^{\infty} k \cdot e^{-\lambda} \frac{\lambda^k}{k!} \text{ first term is } 0.$$

$$= \sum_{k=1}^{\infty} e^{-\lambda} \frac{\lambda^k}{(k-1)!} \text{ cancel the } k.$$

$$= \lambda \sum_{k=1}^{\infty} e^{-\lambda} \frac{\lambda^{k-1}}{(k-1)!} \text{ factor out } \lambda.$$

$$= \lambda \sum_{j=0}^{\infty} e^{-\lambda} \frac{\lambda^j}{(j)!} \text{ Define } j = k - 1$$

$$= \lambda \cdot 1 \text{ The summation is just the pmf!}$$
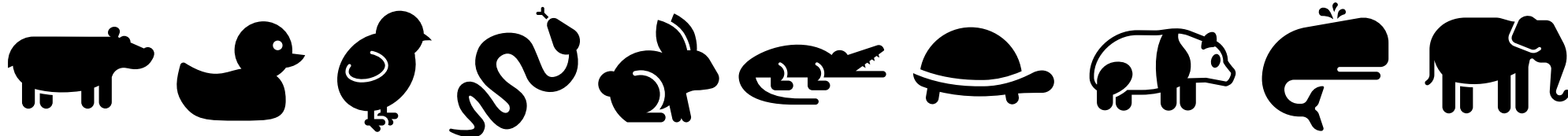
Where did this expression come from?

For the cars we said, "it's like every car in Seattle independently might cause an accident."

If we knew the exact number of cars, and they all had identical probabilities of causing an accident...

It'd be just like counting the number of heads in n flips of a bunch of coins (the coins are just VERY biased).

The Poisson is a certain limit as $n \to \infty$ but np (the expected number of accidents) stays constant.

Zoo!

$X \sim \text{Unif}(a, b)$	$X \sim \text{Ber}(p)$	$X \sim \text{Bin}(n, p)$	$X \sim \text{Geo}(p)$
$p_X(k) = \dfrac{1}{b - a + 1}$ $\mathbb{E}[X] = \dfrac{a + b}{2}$ $\text{Var}(X) = \dfrac{(b - a)(b - a + 2)}{12}$	$p_X(0) = 1 - p;$ $p_X(1) = p$ $\mathbb{E}[X] = p$ $\text{Var}(X) = p(1 - p)$	$p_X(k) = \dbinom{n}{k} p^k (1 - p)^{n-k}$ $\mathbb{E}[X] = np$ $\text{Var}(X) = np(1 - p)$	$p_X(k) = (1 - p)^{k-1} p$ $\mathbb{E}[X] = \dfrac{1}{p}$ $\text{Var}(X) = \dfrac{1 - p}{p^2}$

| $X \sim \text{NegBin}(r, p)$	$X \sim \text{HypGeo}(N, K, n)$	$X \sim \text{Poi}(\lambda)$
$p_X(k) = \dbinom{k - 1}{r - 1} p^r (1 - p)^{k-r}$ $\mathbb{E}[X] = \dfrac{r}{p}$ $\text{Var}(X) = \dfrac{r(1 - p)}{p^2}$	$p_X(k) = \dfrac{\dbinom{K}{k}\dbinom{N-K}{n-k}}{\dbinom{N}{n}}$ $\mathbb{E}[X] = n\dfrac{K}{N}$ $\text{Var}(X) = \dfrac{K(N - K)(N - n)}{N^2(N - 1)}$	$p_X(k) = \dfrac{\lambda^k e^{-\lambda}}{k!}$ $\mathbb{E}[X] = \lambda$ $\text{Var}(X) = \lambda$

Zoo Takeaways

You can do relatively complicated counting/probability calculations much more quickly than you could week 1!

You can now explain why your problem is a zoo variable and save explanation on homework (and save yourself calculations in the future).

Don't spend extra effort memorizing…but be careful when looking up Wikipedia articles.

The exact definitions of the parameters can differ (is a geometric random variable the number of failures before the first success, or the total number of trials including the success?)
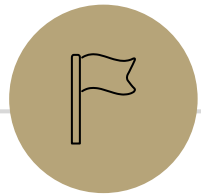
Reviewing what we have learned

What have we done over the past 5 weeks?

Counting

Combinations, permutations, indistinguishable elements, starts and bars, inclusion-exclusion…

Probability foundations

Events, sample space, axioms of probability, expectation, variance

Conditional probability

Conditioning, independence, Bayes' Rule

Refined our intuition

Especially around Bayes' Rule

What's next?

Continuous random variables.

So far our sample spaces have been countable. What happens if we want to choose a random real number?

How do expectation, variance, conditioning, etc. change in this new context?

Mostly analogous to discrete cases, but with integrals instead of sums.

Analysis when it's inconvenient (or impossible) to exactly calculate probabilities.

Central Limit Theorem (approximating discrete distributions with continuous ones)

Tail Bounds/Concentration (arguing it's unlikely that a random variable is far from its expectation)

A first taste of making predictions from data (i.e., a bit of ML)
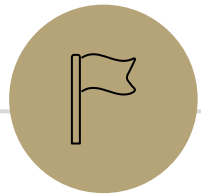
More Practice

Practice: Poisson

Seattle averages 3 days with snowfall per year.

Suppose that the number of days with snow follows a Poisson distribution. What is the probability of getting exactly 5 days of snow?

According to the Poisson model, what is the probability of getting 367 days of snow?

Practice: Poisson

Let $X \sim \text{Poi}(3)$.

$$p_X(5) = \frac{3^5 e^{-3}}{5!} \approx .1008$$

Or about once a decade.

Probability of 367 snowy days, err...

The distribution says

$p_X(367) \approx 1.8 \text{x} 10^{-610}$.

Definition of a "year" says probability should be 0.

Practice Problem: Coin Flips

There are two coins, heads up, on a table in front of you. One is a trick coin – both sides are heads. The other is a fair coin.

You are allowed 2 coin flips (total between the two coins) to figure out which coin is which. What is your strategy? What is the probability of success?

Flip each once

With probability 1 when we flip the trick coin it shows heads.

With probability ½ the fair coin shows tails, and we know it's the fair one.

With probability ½, both the coins were heads and we have learned nothing. So, we have a ½ chance of guessing which is which.

$$\frac{1}{2} \cdot 1 + \frac{1}{2} \cdot \frac{1}{2} = \frac{3}{4}$$ chance of success

Flip one twice.

Now flip the same coin twice.

We'll see a tails with probability $\frac{1}{2} \cdot \frac{3}{4} = \frac{3}{8}$

If we don't see a tails, just guess the other one? What's our probability of guessing right? Let T be the event "we're flipping the trick coin" N be the event we saw no tails

$$\mathbb{P}(T|N) = \frac{\mathbb{P}(N|T)\mathbb{P}(T)}{\mathbb{P}(N)} = \frac{1 \cdot \frac{1}{2}}{1 \cdot \frac{1}{2} + \frac{1}{4} \cdot \frac{1}{2}} = \frac{4}{5}$$

Guess right: $1 \cdot \frac{3}{8} + \frac{4}{5} \cdot \frac{5}{8} = \frac{7}{8}$ Better to flip the same coin twice!

Practice Problem: Donuts

You are buying at most 7 donuts (could be 0, could be 1,..., could be 7).

There are chocolate, strawberry, and vanilla donuts.

How many different orders could you make – give a simple formula!

Donuts: Approach 1

Use the sum rule over the possible numbers of donuts.

For n donuts, by the stars and bars formula there are $\binom{n+3-1}{3-1}$

So we have $\sum_{n=0}^{7}\binom{n+3-1}{3-1}$ correct. But not simple yet...

Use pascal's rule. Rewrite $\binom{2}{2}$ as $\binom{3}{3}$

We'll get $\binom{j}{2} + \binom{j}{3} = \binom{j+1}{3}$, that can combine with $\binom{j+1}{2}$ until you get $\binom{7+3-1+1}{3} = \binom{10}{3}$

Donuts: Approach 2

Clever way: a fourth type of donut: the don't-buy-one donut.

Then we're buying exactly seven donuts of the four types (chocolate, strawberry, vanilla, don't-buy-one)

By stars and bars $\binom{7+4-1}{4-1} = \binom{10}{3}$.