<u>Announcements</u>
- Homework nightmare
- Homework 3 assigned today

STAT 430: Lecture 9

Additional Random Variables

Continuing Chapter 2

*EXAM 1: 2 weeks from

today

<u>October 3rd</u>

Discrete Random Variables

Poisson Random Variables

Discrete Random Variables

Poisson

Poisson Random Variables

Motivation:

What if I have an (almost) infinite number of independent Bernoulli trials and a very small probability of success?

If I know the rate at which a specific outcome will occur is λ occurances per unit, how can I describe the probability of seeing k outcomes?

Statement:

Let X be the number of times a specific outcome is observed during a process if the rate at which it occurs is λ.

Notation:

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

Discrete Random Variables

Poisson

Poisson Random Variables

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

pmf:

The probability of seeing k successes is:

$$p(k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

$$\sum_{k=0}^{\infty} p(k) = 1$$

Cumulative Probability:

$$P(X \leq k) = p(0) + p(1) + p(2) + \ldots + p(k), k = 0, 1, 2, \ldots$$
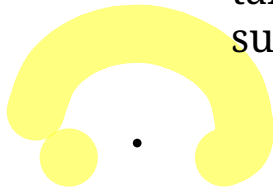
Expectation:

"$E(X)$"

$$\mu = \lambda$$

Variance:

$$\sigma^2 = \lambda$$

Discrete Random Variables

Poisson

$3600 \cdot 24$

$= 72000$

$+ 4 \cdot 7600$

big

Poisson and Binomial Relationship

Consider a binomial distribution for which the number of trials is very high but the probability of a successful outcome is very low. For instance, suppose that we consider our day as being broken into 1 second units of time and watch record a "successful" observation if a message arrives on our phone during that unit. Then we have a massive number of trials but a very low probability that we get a message at any single second.

Now, you may say that even though the probability of getting a message in any single second is low, the probability of getting two messages in a second (though still lower) is not zero - and a **binomial distribution** does not allow two successes on a single trial. That's true - so maybe we shouldn't use a single second as our unit of observation. Maybe we should use 1/1000 of a second. But all we would do is repeat the same argument, only now we are talking about a smaller probability of getting two successes and a smaller unit of time.

Discrete Random Variables

Poisson and Binomial Relationship

Illustration:

Poisson

0   3 100.24
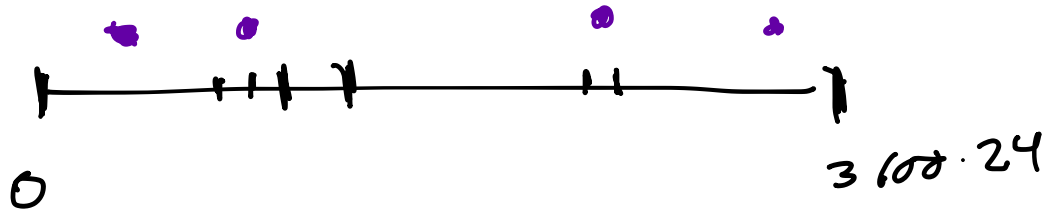
Discrete Random Variables

Poisson

The Poisson distribution can be thought of as the conclusion of this discussion. If we simply say that the intervals will always be a length so that $n \cdot p = \lambda$ then by making the intervals smaller and smaller (i.e., by letting $n \to \infty$) while at the same time keeping $p = \lambda/n \to 0$ we go from this:

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

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

$$= \frac{1}{k!} \frac{n!}{(n-k)!} \left(\frac{\lambda}{n}\right)^k \left(1 - \frac{\lambda}{n}\right)^{n-k}$$

$$= \frac{1}{k!} \lambda^k \left(\frac{n!}{n^k(n-k)!}\right) \left(1 - \frac{\lambda}{n}\right)^{n-k}$$

to

$$p(k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

Discrete Random Variables

Poisson

Poisson and Binomial Relationship

Here are the important take aways:

- For very, very large n and very, very small p we can take $\lambda = n \cdot p$ and approximate the probability of k successful outcomes using a Poisson distribution.
- Distributions based on parameters (the p and the n in this case) may have interesting connections to other distributions by taking limits or fixing the parameters at certain values.

Discrete Random Variables

Hypergeometric and Negative Binomial

Discrete Random Variables

Poisson

Hypergeometric

Hypergeometric Distribution

Suppose that we have n objects r of which are of interest and $n - r$ of which are not of interest. Then if we select m of these objects from the total group and let X be the number of objects we select that were of interest, then X is a **hypergeometric random variable** and the probability that $X = k$ is

$$P(X = k) = \frac{\binom{r}{k}\binom{n-r}{m-k}}{\binom{n}{m}}$$

Example: $\quad n, m, r$
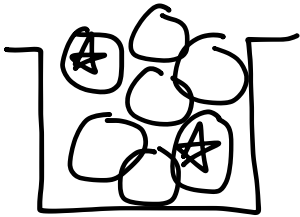
We have an urn with 10 balls with stars on them and 20 balls with stars. If we draw 19 balls from the urn, what is the probability that we will draw 4 balls with stars?

30 balls total $\qquad n = 30, \; r = 10, \; m = 19$

$$P(X = 4) = \frac{\binom{10}{4}\binom{20}{15}}{\binom{30}{19}}$$

Note be careful when $m \geq n - r$ or

$k \geq n - r$

Discrete Random Variables

Poisson

Hypergeometric

Negative Binomial

Negative Binomial Distribution

Suppose that a sequence of independent bernoulli trails, each with the same probability of successs is performed until we have observed the rth successful trial. If we define X as the total number of trials required, r^{th} successful trial then we say that X is a **Negative Binomial random variable** and we can describe the probability that $X = k$ as:

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

Note: If $X_1, X_2, ..., X_r$ are independent geometric random variables with probability of success p, then $Y = X_1 + X_2 + ... + X_r$ is a negative binomial random variable.

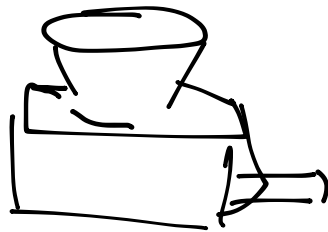# Discrete Random Variables

Poisson

Hypergeometric

Negative Binomial

Negative Binomial Distribution

Example:

A certain expensive piece of equipment needs servicing if the probability that it <u>fails to work</u> on a given trial is at or ~~below~~ <u>0.05</u>. A worker has noticed that the machine has shut down twice during the last 10 uses. If the probability the machine shuts down is 0.05, what is the probability that the user would have seen this result?

$$- - - - - \overset{x}{\underset{10}{-}} - - - \overset{x}{\times}$$

as written:

$$\text{binomial } (10, 0.05) = \frac{10!}{(8!)(2!)} \left(\frac{1}{20}\right)^2 \left(\frac{19}{20}\right)^8$$

Extension

Suppose the only reason the worker is curious about this is because the last trial he saw was a failure.

$$- - - - - | - - - - - \times$$

Negative Binomial: $P(X=10) = \left(\frac{9!}{1! \, 8!}\right) \left(\frac{1}{20}\right)^2 \left(\frac{19}{20}\right)^8$

Continuous Random Variables

Continuous RVs

Background

Continuous Random Variables

We have talked about random values that can take certain specific values - but what if we are interested in a random variable that could take on values in a continuum?

For instance, suppose we are interested in the change in the water level of a lake from yesterday today? Or the weight that a certain type of beam can support?

For these examples, we must create random variables that can mirror the events we are observing - we call these random variables **continuous**

Continuous RVs

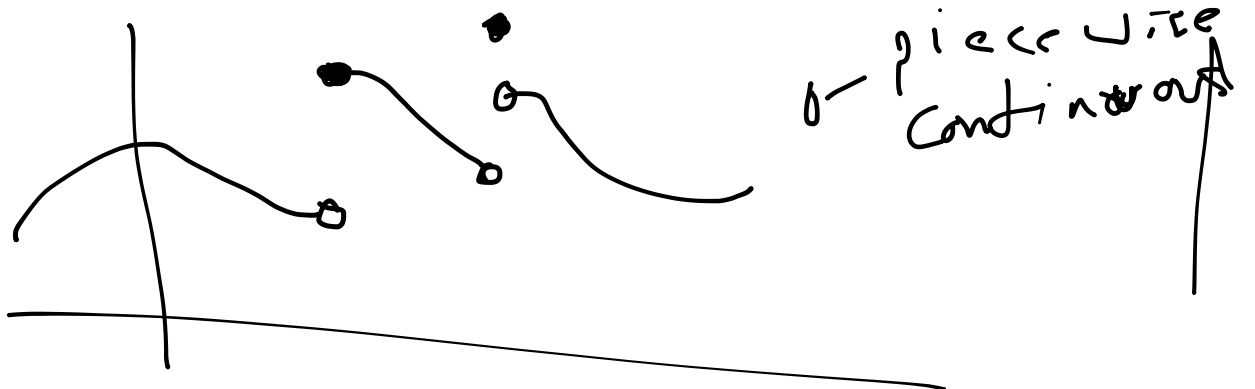## Continuous Random Variables

Background — Probability Density Functions

The main tool that describes how the likelihood is distributed over the possible values is the **probability density function (pdf)**. We use the notation $f(x)$ to describe these functions.

Not any function can be a pdf. A valid pdf will have the following properties:

1. $f(x)$ is non-negative (i.e., $f(x) \geq 0$ for all x)
2. $f(x)$ is piecewise continuous (i.e., it is continuous over open subintervals)
3. The total area under the curve $f(x)$ is 1, i.e., $\int_{-\infty}^{\infty} f(x)dx = 1$.

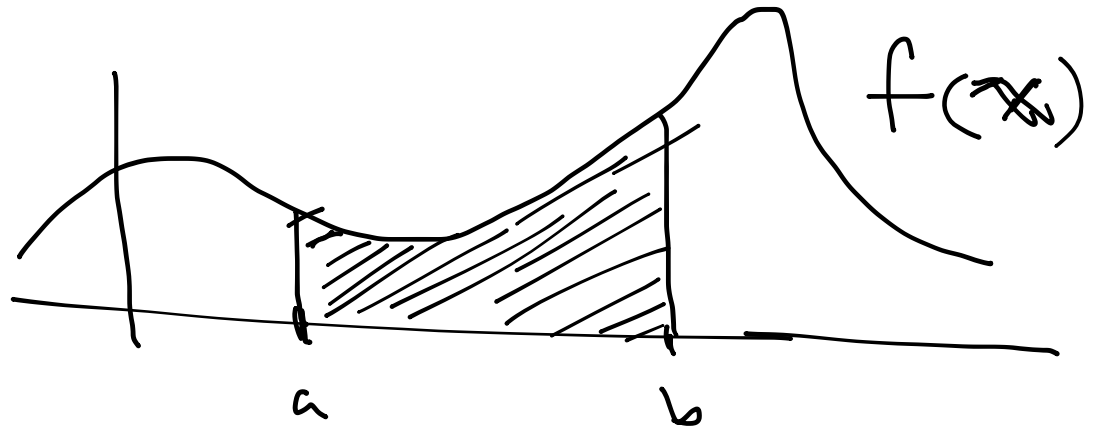The relationship between probabilities and $f(x)$ is described by:

piecewise continuous.

Continuous RVs

Continuous Random Variables

Background

$$P(a < X < b) = \int_a^b f(x)dx$$

$$P(a \leq X \leq a+1) < P(b \leq X \leq b+1)$$

Uniform Random Variables

Motivation:

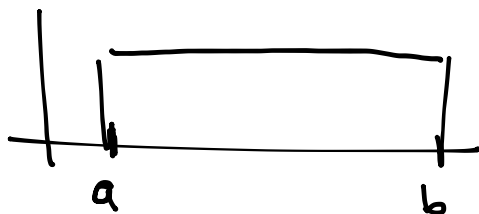I am observing a outcome that must be between two values but the chance that it is in any interval of length α is the same everywhere

Statement: $X \sim Uniform(a, b)$

pdf: $f(x) = \begin{cases} \frac{1}{b-a} & \text{if } a \leq x \leq b \\ 0 & \text{otherwise} \end{cases}$

Expected Value:

$$E(X) = \int_{-\infty}^{\infty} x \cdot f(x) \, dx$$

$$= \int_{a}^{b} x \cdot \frac{1}{b-a} \, dx$$

$$= \frac{1}{b-a} \int_{a}^{b} x \, dx$$

$$= \frac{1}{b-a} \frac{1}{2} x^2 \Big|_{a}^{b} = \frac{1}{2} \frac{b^2 - a^2}{b - a} = \frac{(b-a)(b+a)}{2(b-a)}$$

$$= \frac{b+a}{2}$$