A. **Variance of a Poisson Random Variable.** Let X be a Poisson random variable with parameter λ, for example X could be the number of car crashes on a highway over a week, where λ measures the average. We had calculated:

first moment: $\mathbb{E}[X] = \lambda$

To take advantage of the Taylor series in the margin note, let's calculate:

$\mathbb{E}[X^2 - X] = \mathbb{E}[X(X-1)] =$

second moment: $\mathbb{E}[X^2] =$

> **Poisson Variance.** If $X \sim \text{Poisson}(\lambda)$ then:
>
> $\text{Var}(X) =$

Recall: the variance of X is:

$\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2$

The probability mass function of $\text{Poisson}(\lambda)$ is:

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

if $k \in \mathbb{N}$, and is 0 otherwise.

Recall the Taylor series:

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

and its second derivative:

$e^x = \sum_{k=0}^{\infty} \frac{k(k-1)x^{k-2}}{k!}$

B. **Precursor to Continuous Random Variables.** Let X be the random variable equal to the value of a **uniform** selection from interval $(0,1)$. Then X takes on uncountably many values, so is **not** a discrete random variable.

In the discrete case, uniform applies to **equally likely** selections. This is the idea we are trying to generalize here.

Nonetheless, what does our intution say about:

$\mathbb{P}\left(X \leqslant 1\right) =$ $\qquad$ $\mathbb{P}\left(X \leqslant \frac{1}{2}\right) =$

In general, the **cumulative distribution function** is given by:

So-called because $\mathbb{P}(X \leqslant x)$ calculates the cumulative probability up to the value $X = x$.

$$F(x) = \mathbb{P}\left(X \leqslant x\right) = \begin{cases} & \\ & \end{cases}$$

Let's also consider:

$\mathbb{P}\left(\frac{1}{2} \leqslant X \leqslant \frac{3}{4}\right) =$ $\qquad$ $\mathbb{P}\left(\frac{1}{2} \leqslant X \leqslant \frac{1}{2} + dx\right) =$ $\qquad$ $\mathbb{P}\left(X = \frac{1}{2}\right) =$

This argument can be applied to any individual value to show $\mathbb{P}(X = x) = 0$. So, probabilities of individual values are 0. Therefore, we cannot hope to compute general probabilities by merely adding up probabilities of individual values.

Instead, let us calculate probability over an infinitesimal interval:

$$\mathbb{P}\left(x \leqslant X \leqslant x + dx\right) = \begin{cases} & \\ & \end{cases}$$

We can get the **probability density function** by dividing the probability by the length of the interval.

The units for the probability density function are thus:

$$\frac{\text{probability}}{\text{units of } x}$$

$$f(x) = \frac{\mathbb{P}\left(x \leqslant X \leqslant x + dx\right)}{dx} = \begin{cases} & \\ & \end{cases}$$

In calculus, you learned that integrating density of mass gives you mass. Here, substitute "mass" with "probability" to conclude that integrating the probability density function gives you probability. So, for example:

$\mathbb{P}\left(\frac{1}{4} \leqslant X \leqslant \frac{1}{2}\right) =$

And the cumulative distribution function:

if $x \in (0,1)$, then: $F(x) = \mathbb{P}(X \leqslant x) =$

and therefore, at all points where the cumulative distribution function is differentiable, by the fundamental theorem of calculus, its derivative will equal the probability density function:

$F'(x) =$

C. **Continuous Random Variables.** The example of the random variable given by the value of a uniform selection from the interval $(0, 1)$ is an example of a **continuous random variable**. What characterizes a continuous random variable is that probability of the random variable taking particular values can be found by integrating a **probability density function** over those values.

Continuous Random Variable. A **continuous random variable** on a probability space is a function $X : \Omega \to \mathbb{R}$ that has a **probability density function**.

A **pdf** is a nonnegative "integrable" function $f(x) \geqslant 0$ so that, that for any "measurable" subset D of x–values, we have:

$\mathbb{P}(X \in D) =$

Some consequences include:

zero probability at points: $\mathbb{P}(X = x) =$

total probability: $\int_{-\infty}^{\infty} f(x)\, dx =$

probability density function = **pdf**

"Integrable" and "measurable" are technical terms, that you do not need to know. Intuitively, "integrable" means that the function can be integrated, and "measurable subset" means we can integrate over that subset.

As singular values do not matter, a continuous random variable can have **multiple** pdfs (change the value of the pdf at one value). Instead, when comparing continuous random variables, we use cumulative probability, which **is** unique.

Cumulative Distribution. The **cumulative distribution function** of a discrete or continuouous random variable X is:

$F(x) = \mathbb{P}(X \leqslant x)$

If two random variables X and Y have the same cumulative distribution function, then we say they are **identically distributed** and write $X \sim Y$.

Evidently, the cdf $F(x)$ increases with x, with $F(-\infty) =$ and $F(\infty) =$.

cumulative distribution function = **cdf**

Technically, these are limits as $x \to -\infty$ and $x \to \infty$.

The derivative connection between the cdf and pdf that we observed in the uniform example holds for any continuous random variable.

Cdf and Pdf Connection.

Let X be a **continous** random variable with cdf $F(x)$ and pdf $f(x)$. Then:

$F(x) = \mathbb{P}(X \leqslant x) =$

Therefore $F'(x) \overset{\text{a.e.}}{=}$

This gives us a useful **heuristic**, i.e. informal rule for developing intuition:

$$f(x) = \frac{dF}{dx} \approx \frac{\mathbb{P}(x \leqslant X \leqslant x + dx)}{dx} \longleftarrow \frac{\text{probability}}{\text{length}} \longleftarrow$$

"a.e." stands for "almost everywhere" and is a tehnical term, that you do not need to know. Intuitively, it means that that they have the same integrals over any given "measurable" subset. Practically, it means that you can use $F'(x)$ as a pdf, even if F is not differentiable everywhere.

D. **Continuous Uniform Random Variables.** Let's generalize the example earlier, where we looked at uniform random variable on $(0, 1)$.

Continuous Uniform. Let Ω be the sample space of uniform selections from the interval (a, b), and let X be the value of a selection. Then X is the **uniform random variable** on (a, b), written:

$X \sim \text{Uniform}(a, b)$

and has cumulative distribution function:

$$F(x) = \mathbb{P}(X \leqslant x) = \begin{cases} \\ \\ \end{cases}$$

and has probability density function:

$$f(x) = \begin{cases} \\ \end{cases}$$