2:43 PM Mon Jul 13 40%

MATH 394 B Pr...1 (Summer 2026)

Home Insert Draw View

$\mathbb{P}(X=K) \approx \mathbb{P}(Y=K)$ for any K

Poisson Approximation to Binomial

$X \sim Bin(n,p)$

n & p have certain properties

dist

$\overset{dist}{\approx}$

$Y \sim poisson(\lambda)$

$\lambda = n \times p$

Poisson Approximation to Binomial

When

$$X \sim \text{Bin}(n, p).$$

AND

- n is large;
- p is small;
- $\lambda = np$ is moderate,

$\begin{cases} n \to \infty \\ p \to 0 \\ n \times p \to c \text{ (constant)} \end{cases}$

$p = \frac{1}{n} \Rightarrow \begin{cases} n \to \infty \\ p \to 0 \\ n \times p = n \times \frac{1}{n} = 1 \text{ (constant)} \end{cases}$

Then,

$$\boxed{X \approx \text{Pois}(\lambda).}$$

$\lambda = n \cdot p$

reason? (next slide)

That is,

$$\mathbb{P}(X = k) \overset{def}{=} \underbrace{\binom{n}{k}}_{\to \frac{n!}{k!(n-k)!}} p^k (1-p)^{n-k} \approx e^{-\lambda} \frac{\lambda^k}{k!}.$$

PMF of poisson$(\lambda = n \cdot p)$

$\rightarrow \frac{n!}{K!(n-K)!}$ PMF of

Why the Poisson Approximation Works

Assume $X \sim \text{Bin}(n, p)$ with $p = \lambda/n$ ($\rightarrow 0$) where $\lambda > 0$ is fixed. For each fixed $k = 0, 1, 2, \ldots$, we have

$$\frac{n(n-1)\cdots(n-k+1)}{n \cdot n \cdots n}$$

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

$$\mathbb{P}_{Binom(n,p)}(X = k) = \binom{n}{k}\left(\frac{\lambda}{n}\right)^k\left(1 - \frac{\lambda}{n}\right)^{n-k} = \frac{n!}{k!(n-k)!\,n^k}\lambda^k\left(1 - \frac{\lambda}{n}\right)^n\left(1 - \frac{\lambda}{n}\right)^{-k}$$

$p = \lambda/n$

$$= \underbrace{\left[(1)\left(1 - \frac{1}{n}\right)\left(1 - \frac{2}{n}\right)\cdots\left(1 - \frac{k-1}{n}\right)\right]}_{\xrightarrow[n\to\infty]{} 1}\underbrace{\frac{\lambda^k}{k!}}_{\text{stays fixed}}\underbrace{\left(1 - \frac{\lambda}{n}\right)^n}_{\xrightarrow[n\to\infty]{} e^{-\lambda}}\underbrace{\left(1 - \frac{\lambda}{n}\right)^{-k}}_{\xrightarrow[n\to\infty]{} 1}$$

$1^{-k} = 1$

example) $\lim \left(1 + \frac{1}{n}\right)^n = e$

$$\frac{n!}{(n-k)!} = n\cdots(n-k+1)$$

$$\longrightarrow 1 \cdot \frac{\lambda^k}{k!} \cdot e^{-\lambda} \cdot 1 = e^{-\lambda}\frac{\lambda^k}{k!} = \mathbb{P}_{X\sim Poisson(\lambda)}(X = k)$$

$$\lim_{n\to\infty}(1 + a_n)^{b_n} = e^{\lim_{n\to\infty} a_n \cdot b_n}$$

if $\lim a_n \cdot b_n < \infty$ & $\lim a_n^2 \cdot b_n = 0$

Thus,

$$\boxed{\mathbb{P}_{X\sim Bin(n,p)}(X = k) \xrightarrow{n\to\infty} e^{-\lambda}\frac{\lambda^k}{k!}.}$$

This phenomenon is formally known as convergence in distribution (or weak convergence) in Probability theory. We will see more with Central Limit Theorem (CLT) and Law of Large Numbers (LLN) in Chapter 10B (Limit Theorems).

Weak convergence

Visualization of Poisson Approximation to the Binomial

Poisson Approximation to the Binomial

As n increases and p decreases with $\lambda = np$ fixed, $\mathrm{Bin}(n, p) \approx \mathrm{Pois}(\lambda)$

$P(X=3) = \binom{200}{3}\left(\frac{1}{200}\right)^{3}$

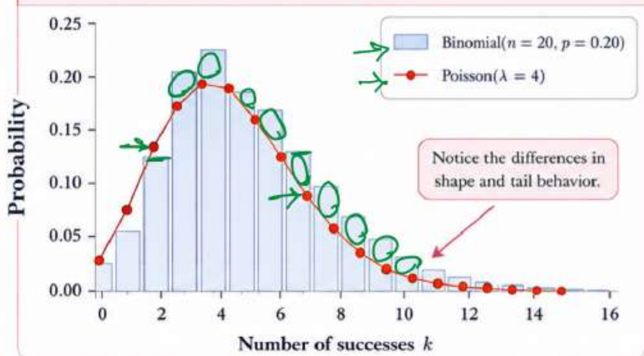
$X \sim Bin(2000, \frac{1}{2000})$

Here we keep $\lambda = np = 4$ fixed.

1) $n = 20, \quad p = 0.20, \quad \lambda = np = 4$

Approximation is coarse

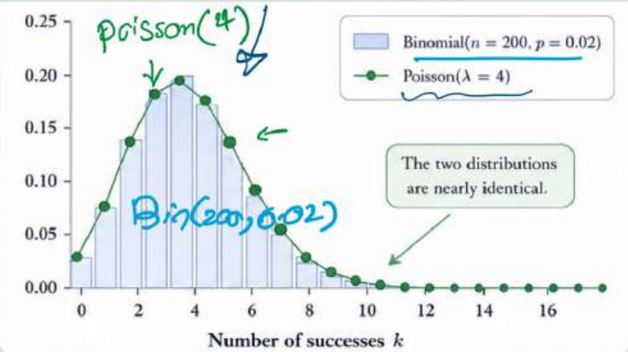
Binomial and Poisson are visibly different.



2) $n = 200, \quad p = 0.02, \quad \lambda = np = 4$

Approximation is very good

Binomial and Poisson are almost indistinguishable.



Takeaway

When events are rare (small p) and trials are many (large n), while the mean $\lambda = np$ stays fixed, the Binomial distribution is well approximated by a Poisson distribution with mean λ.

Example: Website Visitors

Suppose ($n=$) 1,000,000 people independently decide whether to visit a website.
Each visits with probability

$$\to \quad p = 2 \times 10^{-6}. \qquad n = 10^6$$

Let

$$X = \text{number of visitors}. \qquad n \times p = 2$$

Then

$$X \sim \text{Bin}(1{,}000{,}000, 2 \times 10^{-6}). \to P(X=10) = \binom{10^6}{10}\left(\frac{2}{10^6}\right)^{10}\left(1-\frac{2}{10^6}\right)^{10^6-10}$$

Here

$$\to \lambda = np = 2.$$

So

$$X \approx \text{Pois}(2). \to \text{PMF: } \frac{e^{-\lambda}\cdot\lambda^{x}}{x!}$$

For example,

$$\mathbb{P}(X = 0) \approx e^{-2}.$$

Note that it is computationally hard to compute directly from the Binomial PMF:

$$\mathbb{P}(X = k) = \frac{1000000!}{k!(1000000 - k)!}\left(\frac{2}{10^6}\right)^k\left(1 - \frac{2}{10^6}\right)^{1000000-k} = \tag{4}$$

Summary

- Expectation is a probability-weighted average:

$$\mathbb{E}(X) = \sum_x x\mathbb{P}(X = x).$$

- Linearity of Expectation

$$\mathbb{E}(X + Y) = \mathbb{E}(X) + \mathbb{E}(Y),$$

  with no independence required.
- Fundamental bridge:

$$\mathbb{E}(I_A) = \mathbb{P}(A).$$

- Variance measures spread: (not linear unless X & Y are indep.)

$$\mathrm{Var}(X) = \mathbb{E}(X^2) - (\mathbb{E}X)^2.$$

- Poisson approximates Binomial when n is large, p is small, and $\lambda = np$ is moderate.

Next lecture

Continuous random variables.

chapter5a_continuous_uniform

Monday, July 13, 2026 12:23 PM

chapter5a_continuous...

Chapter 5A: Continuous Random Variables, PDFs, and Uniform Distribution
MATH/STAT 394: Probability I

Arman Jahangiri

University of Washington Department of Mathematics

Summer 2026

Learning Goals

By the end of this lecture, students should be able to:

- define and interpret PDFs and CDFs;
- compute probabilities by integrating a PDF;
- recognize valid PDFs;
- compute expectation for continuous random variables;
- define and use the Uniform distribution;
- understand the universality of the Uniform distribution.

Continuous Random Variables

From Discrete to Continuous Random Variables

Discrete random variables take values in a finite or countably infinite set.

Continuous random variables can take any value in an interval, such as

$$(0,1), \quad (a,b), \quad [0,\infty), \quad \mathbb{R}.$$

0.1 $\quad \frac{1}{n} < 0.1$

Main difference

For a continuous random variable,
$$\boxed{\mathbb{P}(X = x) = 0}$$
for every individual point x.

$x-\varepsilon \quad x \quad x+\varepsilon$

By contradiction, suppose $\mathbb{P}(X=x) = a_x > 0$. Then for any x, $\exists n \in \mathbb{N}$ s.th. $\left(\frac{1}{n_x} < a_x\right)$

$1 = P(S) = \sum_x P(X=x) = \sum_x a_x > \sum_{x \text{ uncountable}} \frac{1}{n_x} \geq \sum_{n=1}^{\infty} \frac{1}{n} = \infty$ (countable set) $\times$.

Probability Density Functions

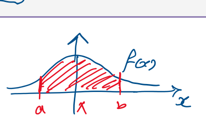
Similar to PMFs.

Definition

For a continuous random variable X, the probability density function, or PDF, is a function f such that
$$\mathbb{P}(X \in A) = \int_A f(x)\,dx, \quad \forall A \subseteq \mathbb{R}$$

The value $f(x)$ itself is not a probability (could be greater than 1).
Probabilities are areas under the density (pdf) curve.

$$p(X \in A) = \sum_{x \in A} p(x)$$



How to verify a function is a PDF?

A function f is a valid PDF if it satisfies two properties, namely
$$\boxed{f(x) \geq 0} \quad \text{for all } x,$$
and
$$\boxed{\int_{-\infty}^{\infty} f(x)\,dx = 1.}$$

Analogy

For discrete random variables:
$$p_X(x) \geq 0, \quad \sum_x p_X(x) = 1.$$

For continuous random variables:
$$f_X(x) \geq 0, \quad \int_{-\infty}^{\infty} f(x)\,dx = 1.$$

CDF and PDF

CDF has the same formula for continuous and discrete r.v.s, i.e.

(2) $\boxed{P(X \in A) = \int_A f(x)\,dx}$

$$F(x) = \mathbb{P}(X \leq x). = \begin{cases} \sum_{k \leq x} P(k) & X\text{: Discrete} \\ \int_{-\infty}^{x} f(t)\,dt & X\text{: cont.} \end{cases} \quad \text{by (1) \& (2)}$$

For a continuous random variable with PDF f, according to the definition of PDF,

continuous $\quad F(x) = \int_{-\infty}^{x} f(t)\,dt. \quad A = (-\infty, x]$

$P(X \leq x) = P(X \in (-\infty, x])$ (1) A

Note that , according to the definition of CDF,
$$(*) \quad \boxed{\mathbb{P}(a < X \leq b) = F(b) - F(a)}$$

$\{a < X \leq b\} = \{X \leq b\} - \{X \leq a\}$

This, together with the Fundamental Theorem of Calculus, implies that
$$\frac{d}{dx}F(x) = F'(x) = f(x) \tag{1}$$

For **continuous** r.v.s, since $\mathbb{P}(X = a) = 0$ for any $a \in \mathbb{R}$,
$$\mathbb{P}(X \leq a) = \mathbb{P}(X < a),$$
$$\mathbb{P}(X \geq a) = \mathbb{P}(X > a),$$
$$\mathbb{P}(a \leq X \leq b) = \mathbb{P}(a < X \leq b) = \mathbb{P}(a \leq X < b) = \mathbb{P}(a < X < b) = F(b) - F(a)$$

calculus: if $g(x) = \int_a^x f(t)\,dt \Rightarrow \boxed{\frac{d}{dx} g(x) = f(x)}$

$\{X \leq a\} = \{X < a\} \cup \{X = a\} \Rightarrow \boxed{P(X \leq a) = P(X < a) + \cancel{P(X = a)}}$

What Does the Density $f(x)$ Mean?

For a continuous random variable,
$$F(x) = \mathbb{P}(X \leq x) = \int_{-\infty}^{x} f(t)\,dt.$$

For a small number $\boxed{\epsilon > 0}$
$$\mathbb{P}(x \leq X \leq x + \epsilon) = F(x + \epsilon) - F(x).$$

small $\quad x \quad x+\varepsilon$

Therefore,
$$\frac{\mathbb{P}(x \leq X \leq x + \epsilon)}{\epsilon} = \boxed{\frac{F(x + \epsilon) - F(x)}{\epsilon}}.$$

$p(X \in [x, x+\varepsilon]) = ?$

Taking the limit as $\epsilon \to 0$,
$$\lim_{\epsilon \to 0} \frac{\mathbb{P}(x \leq X \leq x + \epsilon)}{\epsilon} = \lim_{\epsilon \to 0} \frac{F(x + \epsilon) - F(x)}{\epsilon}.$$

$f: \mathbb{R} \to \mathbb{R}$

By the definition of derivative,
$$f(x) = F'(x) = \boxed{\lim_{\epsilon \to 0} \frac{F(x + \epsilon) - F(x)}{\epsilon}}.$$

$\frac{d}{dx} F(x_0) = \lim_{\varepsilon \to 0} \frac{F(x_0+\varepsilon) - F(x_0)}{\varepsilon}$

$F'(x_0) = f(x_0)$

Thus,
$$f(x) = \lim_{\epsilon \to 0} \frac{\mathbb{P}(x \leq X \leq x + \epsilon)}{\epsilon}$$

$\varepsilon f(x) \approx P(x \leq X \leq x+\varepsilon)$ for small $\varepsilon > 0$

Expectation for Continuous Random Variables

Definition

If X is continuous with PDF f, then
$$\boxed{\mathbb{E}(X) = \int_{-\infty}^{\infty} x f(x)\,dx.}$$

If the support is an interval, we integrate over that interval.

Expectation is still the balancing point of the distribution.

Continuous LOTUS

LOTUS

If X has PDF f and $g : \mathbb{R} \to \mathbb{R}$, then
$$\boxed{\mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(x) f(x)\,dx.}$$

Proof idea. For a small interval $[x, x + \Delta x]$,
$$\mathbb{P}(x \leq X \leq x + \Delta x) \approx f(x)\Delta x.$$
On this interval, $g(X)$ is approximately $g(x)$. Hence,
$$g(X)\mathbb{P}(x \leq X \leq x + \Delta x) \approx g(x) f(x)\Delta x.$$
Adding over many small intervals,
$$\mathbb{E}(g(X)) \approx \sum_i g(x_i) f(x_i)\Delta x.$$
Taking the limit as $\Delta x \to 0$, the Riemann sum becomes
$$\mathbb{E}(g(X)) = \int_{-\infty}^{\infty} g(x) f(x)\,dx.$$

Uniform Distribution

Uniform Distribution

Definition

A random variable U has the Uniform distribution on (a, b) if its PDF is
$$f(x) = \begin{cases} \frac{1}{b-a}, & \boxed{a < x < b,} \\ 0, & \text{otherwise.} \end{cases}$$

$\left\{ \text{Disc Unif}(1, \dots, n) \quad p(X = k) = \frac{1}{n} \right.$

Unif(0,1)

$P(0 < X < \tfrac{1}{2}) = \int_0^{\frac{1}{2}} \frac{1}{1-0}\,dx = \int_0^{\frac{1}{2}} \frac{1}{1-0}\,dx = \frac{1}{2}$

We write
$$U \sim \text{Unif}(a, b).$$

(uniform (0,1)) $p(X \in A) = \text{length}(A)$

$a \leq b \quad (a, b)$

PDF Verification

$P(\tfrac{1}{2} < X < 1) = \int_{\frac{1}{2}}^{1} \frac{1}{1-0}\,dx = 1 - \frac{1}{2} = \frac{1}{2}$

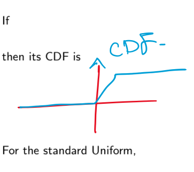
$$f(x) = \frac{1}{b-a} > 0$$
since $b > a$. And
$$\int_{-\infty}^{\infty} f(x)\,dx = \boxed{\int_a^b \frac{1}{b-a}\,dx} = \frac{1}{b-a}(b-a) = 1.$$

$\int_a^b \frac{1}{b-a}\,dx = \frac{1}{b-a} \cdot \int_a^b dx = \frac{1}{b-a}\left(x\Big]_a^b\right) = \frac{b-a}{b-a} = 1$

$\boxed{\int_{-\infty}^{\infty} f(x)\,dx} = \underbrace{\int_{-\infty}^{a} 0\,dx}_{0} + \int_a^b \frac{1}{b-a}\,dx + \underbrace{\int_b^{\infty} 0\,dx}_{0}$

Uniform CDF

If
$$U \sim \text{Unif}(a, b),$$
then its CDF is
$$F(x) = \begin{cases} 0, & x \leq a, \\ \frac{x-a}{b-a}, & a < x < b, \\ 1, & x \geq b. \end{cases}$$



$F(x) = P(X \leq x) = \int_a^x f(t)\,dt = \int_a^x \frac{1}{b-a}\,dt = \frac{1}{b-a}\int_a^x 1\,dt = \frac{1}{b-a} \cdot \left[t\Big|_a^x\right] = \frac{x-a}{b-a}$

For the standard Uniform,
$$U \sim \text{Unif}(0, 1),$$
we have
$$F(x) = x, \qquad 0 < x < 1.$$

Universality of the Uniform

Why the Uniform Distribution Is Special

The Uniform distribution is a universal starting point for simulation.

Many algorithms first generate
$$U \sim \text{Unif}(0, 1),$$
then transform U to obtain another distribution.

Main idea

A Uniform random variable can be transformed into a random variable with many other distributions.

Inverse CDF Method

Suppose F is a continuous, strictly increasing CDF.
Let
$$U \sim \text{Unif}(0, 1).$$
Define
$$X = F^{-1}(U).$$
Then
$$X \text{ has CDF } F.$$

This is called the inverse CDF method or inverse transform sampling.

Why the Inverse CDF Method Works

Let
$$X = F^{-1}(U).$$
Then
$$\mathbb{P}(X \leq x) = \mathbb{P}(F^{-1}(U) \leq x).$$
Since F is increasing,
$$F^{-1}(U) \leq x \iff U \leq F(x).$$
Therefore,
$$\mathbb{P}(X \leq x) = \mathbb{P}(U \leq F(x)).$$
Because $U \sim \text{Unif}(0, 1)$,
$$\mathbb{P}(U \leq F(x)) = F(x).$$
Thus X has CDF F.

Probability Integral Transform

Conversely, if X has continuous strictly increasing CDF F, then
$$F(X) \sim \text{Unif}(0, 1).$$

Interpretation

Applying the CDF to a random variable converts its value into a percentile.

For example, if exam scores have CDF F, then $F(X)$ is the percentile of a randomly selected student.

Uniform Mean and Variance

If $U \sim \text{Unif}(a, b)$, then
$$f_U(u) = \frac{1}{b-a}, \qquad a \leq u \leq b.$$

Mean

$$\mathbb{E}(U) = \int_a^b u \frac{1}{b-a}\,du = \frac{1}{b-a}\left[\frac{u^2}{2}\right]_a^b = \frac{b^2 - a^2}{2(b-a)} = \frac{a+b}{2}.$$

Variance

First,
$$\mathbb{E}(U^2) = \int_a^b u^2 \frac{1}{b-a}\,du = \frac{b^3 - a^3}{3(b-a)} = \frac{a^2 + ab + b^2}{3}.$$
Thus,
$$\text{Var}(U) = \mathbb{E}(U^2) - [\mathbb{E}(U)]^2 = \frac{a^2 + ab + b^2}{3} - \left(\frac{a+b}{2}\right)^2 = \frac{(b-a)^2}{12}.$$

Location-Scale Transformation

If
$$U \sim \text{Unif}(0, 1),$$
then
$$a + (b-a)U \sim \text{Unif}(a, b).$$
This transforms:
$$(0, 1) \longrightarrow (a, b).$$

Useful idea

Start with the standard Uniform and shift/scale to get any Uniform interval.

Summary

- Continuous random variables have probabilities over intervals, not points.
- A PDF integrates to probabilities:
$$\mathbb{P}(a < X < b) = \int_a^b f(x)\,dx.$$
- A valid PDF is nonnegative and integrates to 1.
- The CDF is accumulated area:
$$F(x) = \int_{-\infty}^{x} f(t)\,dt.$$
- Uniform probabilities are proportional to interval length.
- The Uniform distribution is universal through the inverse CDF method.

Next lecture

Normal and Exponential distributions.