Lecture 17: The Normal Distribution

Slide Credit: Based on Stefano Tessaro’s slides for 312 19au incorporating ideas from Anna Karlin, Alex Tsun, Rachel Lin, Hunter Schafer & myself 😊
Agenda

• Law of Total Expectation (LTE) Practice
• Chebyshev’s Inequality
• The Normal Distribution
• Practice with Normals
Example: Flipping Coins

Suppose wanted to analyze flipping a random number of coins. Suppose someone gave us $Y \sim Poi(5)$ fair coins and we wanted to compute the expected number of heads $X$ from flipping those coins.
Agenda

• Law of Total Expectation (LTE) Practice
• Chebyshev’s Inequality
• The Normal Distribution
• Practice with Normals
Using variance

• If we know more about the random variable, e.g. its variance, we can get a better bound!
Chebyshev’s Inequality

**Theorem.** Let $X$ be a random variable. Then, for any $t > 0$,

$$
P(|X - \mathbb{E}(X)| \geq t) \leq \frac{\text{Var}(X)}{t^2}.
$$

**Proof:** Define $Z = X - \mathbb{E}(X)$

Markov’s inequality ($Z^2 \geq 0$)

Definition of Variance

$$
P(|Z| \geq t) = P(Z^2 \geq t^2) \leq \frac{\mathbb{E}(Z^2)}{t^2} = \frac{\text{Var}(X)}{t^2}.
$$
Let $X$ be Binomial RV with parameters $n, p = 0.5$

$$\mathbb{E}(X) = \frac{n}{2} \quad Var(X) =$$

What is the probability that $X \geq \frac{3n}{4}$?

Chebychev’s inequality: $\mathbb{P} \left( X \geq \frac{3n}{4} \right) \leq$

Markov’s inequality: $\mathbb{P} \left( X \geq \frac{3n}{4} \right) \leq \frac{4}{3n} \cdot \frac{n}{2} = \frac{2}{3}$
Tail Bounds

Useful for approximations of complex systems. How good the approximation is depends on the actual distribution and the context you are using it in.

– Usually loose upper-bounds are okay when designing for worst-case

Generally, the more you know about your random variable the better tail bounds you can get.
Agenda

• Law of Total Expectation (LTE) Practice
• Chebyshev’s Inequality
• The Normal Distribution
• Practice with Normals
Review – Continuous RVs

Probability Density Function (PDF).
\[ f: \mathbb{R} \to \mathbb{R} \text{ s.t.} \]
- \[ f(x) \geq 0 \text{ for all } x \in \mathbb{R} \]
- \[ \int_{-\infty}^{+\infty} f(x) \, dx = 1 \]

Cumulative Density Function (CDF).
\[ F(y) = \int_{-\infty}^{y} f(x) \, dx \]

Theorem. \[ f(x) = \frac{dF(x)}{dx} \]

Density ≠ Probability!

\[ F(y) = \mathbb{P}(X \leq y) \]
Review – Continuous RVs

\[ f_X(x) \]

\[ \mathbb{P}(X \in [a, b]) = \int_a^b f_X(x) \, dx = F_X(b) - F_X(a) \]
**Exponential Distribution**

**Definition.** An exponential random variable $X$ with parameter $\lambda \geq 0$ is follows the exponential density

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

We write $X \sim \text{Exp}(\lambda)$ and say $X$ that follows the exponential distribution.

**CDF:** For $y \geq 0$,

$$F_X(y) = 1 - e^{-\lambda y}$$

The graph illustrates the CDF for different values of $\lambda$: $\lambda = 2$, $\lambda = 1.5$, $\lambda = 1$, and $\lambda = 0.5$. The curves show how the probability accumulates as $y$ increases for each value of $\lambda$. The higher the value of $\lambda$, the steeper the decline in the probability as $y$ increases.
Definition. A Gaussian (or normal) random variable with parameters $\mu \in \mathbb{R}$ and $\sigma \geq 0$ has density

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

(We say that $X$ follows the Normal Distribution, and write $X \sim \mathcal{N}(\mu, \sigma^2)$)
The Normal Distribution

**Definition.** A Gaussian (or normal) random variable with parameters $\mu \in \mathbb{R}$ and $\sigma \geq 0$ has density

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

(We say that $X$ follows the Normal Distribution, and write $X \sim \mathcal{N}(\mu, \sigma^2)$)

**Fact.** If $X \sim \mathcal{N}(\mu, \sigma^2)$, then $\mathbb{E}(X) = \mu$, and $\text{Var}(X) = \sigma^2$

Expectation follows from density being symmetric around $\mu$, $f_X(\mu - x) = f_X(\mu + x)$
The Normal Distribution

Aka a “Bell Curve” (imprecise name)

\[ \mu = 0, \sigma^2 = 3 \]

\[ \mu = -7, \sigma^2 = 6 \]

\[ \mu = 0, \sigma^2 = 1 \]
**Fact.** If \( X \sim \mathcal{N}(\mu, \sigma^2) \), then \( Y = aX + b \sim \mathcal{N}(a\mu + b, a^2 \sigma^2) \)

**Proof.**

\[
\begin{align*}
\mathbb{E}(Y) &= a \mathbb{E}(X) + b = a\mu + b \\
\text{Var}(Y) &= a^2 \text{Var}(X) = a^2 \sigma^2
\end{align*}
\]

Can show with algebra that the PDF of \( Y = aX + b \) is still normal.

Note: \( \frac{X-\mu}{\sigma} \sim \mathcal{N}(0, 1) \)
CDF of normal distribution

**Fact.** If $X \sim \mathcal{N}(\mu, \sigma^2)$, then $Y = aX + b \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$

**Standard (unit) normal** $Z \sim \mathcal{N}(0, 1)$

**CDF.** $\Phi(z) = \mathbb{P}(Z \leq z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-x^2/2} \, dx$ for $Z \sim \mathcal{N}(0, 1)$

Note: $\Phi(z)$ has no closed form – generally given via tables
Table of $\Phi(z)$ CDF of Standard Normal Distn

Make sure to use the one linked on the site!
Example

Let $X \sim \mathcal{N}(0.4, 4)$.

$\mathbb{P}(X \leq 1.2)$
Example

Let $X \sim \mathcal{N}(0.4, 4 = 2^2)$. 

$$
\mathbb{P}(X \leq 1.2) = \mathbb{P}\left( \frac{X - 0.4}{2} \leq \frac{1.2 - 0.4}{2} \right) 
\approx \Phi(0.4) \approx 0.6554
$$
Example

Let $X \sim \mathcal{N}(3, 16)$.

$\mathbb{P}(2 < X < 5)$
Example

Let $X \sim \mathcal{N}(3, 16)$.

\[
P(2 < X < 5) = P\left(\frac{2 - 3}{4} < \frac{X - 3}{4} < \frac{5 - 3}{4}\right)
\]

\[
= P\left(-\frac{1}{4} < Z < \frac{1}{2}\right)
\]

\[
= \Phi\left(\frac{1}{2}\right) - \Phi\left(-\frac{1}{4}\right)
\]

\[
= \Phi\left(\frac{1}{2}\right) - \left(1 - \Phi\left(\frac{1}{4}\right)\right) \approx 0.29017
\]
Example – Off by Standard Deviations

Let $X \sim \mathcal{N}(\mu, \sigma^2)$.

$$
\mathbb{P}(|X - \mu| < k\sigma) =
$$
Example – Off by Standard Deviations

Let \( X \sim \mathcal{N}(\mu, \sigma^2) \).

\[
\mathbb{P}(|X - \mu| < k\sigma) = \mathbb{P}\left(\frac{|X - \mu|}{\sigma} < k\right) = \\
= \mathbb{P}\left(-k < \frac{X - \mu}{\sigma} < k\right) = \Phi(k) - \Phi(-k)
\]
e.g. \( k = 1: 68\%, k = 2: 95\%, k = 3: 99\% \)
Summary of procedure for doing calculations with normal r.v.

If $X \sim \mathcal{N}(\mu, \sigma^2)$, then $\frac{X - \mu}{\sigma} \sim \mathcal{N}(0, 1)$

Therefore,

$$F_X(z) = \mathbb{P}(X \leq z) = \mathbb{P}\left(\frac{X - \mu}{\sigma} \leq \frac{z - \mu}{\sigma}\right) = \Phi\left(\frac{z - \mu}{\sigma}\right)$$
CDF of normal distribution

Fact. If $X \sim \mathcal{N}(\mu, \sigma^2)$, then $Y = aX + b \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$

Standard (unit) normal $Z \sim \mathcal{N}(0, 1)$

CDF. $\Phi(z) = \mathbb{P}(Z \leq z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-x^2/2} \, dx$ for $Z \sim \mathcal{N}(0, 1)$

Note: $\Phi(z)$ has no closed form – generally given via tables

If $X \sim \mathcal{N}(\mu, \sigma^2)$, then $F_X(z) = \mathbb{P}(X \leq z) = \mathbb{P}\left(\frac{X-\mu}{\sigma} \leq \frac{z-\mu}{\sigma}\right) = \Phi\left(\frac{z-\mu}{\sigma}\right)$
Closure of the normal -- under addition

**Fact.** If $X \sim \mathcal{N}(\mu_X, \sigma_X^2)$, $Y \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$ (both independent normal RV) then $aX + bY + c \sim \mathcal{N}(a\mu_X + b\mu_Y + c, a^2\sigma_X^2 + b^2\sigma_Y^2)$

Note: The special thing is that the sum of normal RVs is still a normal RV.

The values of the expectation and variance is not surprising.

- Linearity of expectation (always true)
- When $X$ and $Y$ are independent, $Var(aX + bY) = a^2Var(X) + b^2Var(Y)$