Summary of Last Lecture

- **Markov's inequality** (for non-neg. r.v.'s)
  \[ \Pr \{ X \geq c \} \leq \frac{1}{c} E[X] \]

- **Chebyshev's inequality** (for all r.v.'s)
  \[ \Pr \{ |X - E[X]| \geq c \} \leq \frac{1}{c^2} \text{Var}(X) \]
  \[ \Pr \{ |X - E[X]| \geq c \sigma(X) \} \leq \frac{1}{c^2} \]
Summary of Last Lecture (cont.)

- **Statistical estimation:**
  
  \[ X_1, X_2, \ldots, X_N \] are i.i.d. r.v.'s with expectation \( E[X_i] = \mu \), variance \( \text{Var}(X_i) = \sigma^2 \).

  Estimate of \( \mu \) is: \( \hat{\mu} = \frac{1}{N} (X_1 + \ldots + X_N) \)

  **Thm:** If we take \( N \geq \frac{\sigma^2}{\mu^2} \cdot \frac{1}{\varepsilon^2} \) samples, then

  \[
  \Pr \left[ \left| \hat{\mu} - \mu \right| \geq \varepsilon \mu \right] \leq \delta
  \]

- This is (a quantitative version of) the **Law of Large Numbers**
Continuous Probability

Up to now all our probability spaces were discrete i.e., finite or countably infinite

- Specify $\Pr[\omega]$ for each $\omega \in \Omega$
- $0 \leq \Pr[\omega] \leq 1$
- $\sum_{\omega \in \Omega} \Pr[\omega] = 1$

Note: This implies all random variables are also discrete (i.e., take on at most countably many values, e.g., 0, 1, 2, 3, ... )
What if our prob. space is uncountable?

E.g. “wheel of fortune”

Pointer can end up at any position in \([0, \ell]\), where 
\(\ell = \) circumference of wheel
(or, equivalently, at any angle in \([0, 2\pi]\) \(\rightarrow\) uncountably many outcomes

Compare roulette wheel: only 38 outcomes
How do we assign probabilities to outcomes?

- For each \( \omega \in [0,1] \),
  \( \Pr[\omega] = ?? \)
- \( \sum_{\omega \in [0,1]} \Pr[\omega] = 1 \) ??

**Solution**: Instead assign probabilities to intervals:

for \( 0 \leq a < b \leq 1 \),

\[
\Pr[[a,b]] = \frac{\text{length of } [a,b]}{\text{length of } [0,1]} = \frac{b-a}{l}
\]
Solution: Instead assign probabilities to intervals:

for \(0 \leq a < b \leq l\),

\[
\Pr([a, b]) = \frac{\text{length of } [a, b]}{\text{length of } [0, l]} = \frac{b - a}{l}
\]

These intervals are now our atomic/basic events (replacing sample points \(\omega\) before).

Note that \(\Pr([0, l]) = 1\) and \(\Pr(a) = \Pr([a, a]) = 0\).

We can then compute the probability of any event that can be expressed in terms of intervals — e.g. \(\Pr(U I_i) = \sum_i \Pr[I_i]\) for disjoint intervals \(I_i\).

General theory of continuous probability spaces \(\longrightarrow\) measure theory.
Continuous Random Variables

E.g. let \( X = \) position of pointer in wheel of fortune

Range of \( X \) is the continuous interval \([0, l]\)

Again, \( \Pr [X = a] = 0 \quad \forall a \)

But we can define \( \Pr [a \leq X \leq b] = \frac{b-a}{l} \)

To make this more general, we need the idea of probability density
Definition: A probability density function (p.d.f.) for a continuous r.v. $X$ is a function $f : \mathbb{R} \to \mathbb{R}$ satisfying:

- $f(x) \geq 0 \quad \forall x \in \mathbb{R}$

- $\int_{-\infty}^{\infty} f(x) \, dx = 1$

Then the distribution of $X$ is defined by

$$ \Pr [a \leq X \leq b] = \int_{a}^{b} f(x) \, dx \quad \forall a < b $$

Total area under $f(x)$

$$ = \int_{-\infty}^{\infty} f(x) \, dx = 1 $$
**Example:** Wheel of fortune

Here $X$ is uniform on $[0, l]$, i.e., $\Pr[a \leq X \leq b] = \frac{b-a}{l}$

**P.d.f.:**

\[
f(x) = \begin{cases} 
0 & x < 0 \\
0 & x > l \\
c & 0 \leq x \leq l 
\end{cases}
\]

\[
\int_{-\infty}^{\infty} f(x) \, dx = c l = 1 \quad \Rightarrow \quad c = \frac{1}{l}
\]

For $0 \leq a < b \leq l$:

\[
\Pr[a \leq X \leq b] = \int_{a}^{b} f(x) \, dx = c x \bigg|_{a}^{b} = \frac{b-a}{l}
\]
Comparison with discrete distributions

**Histogram**

\[ P(a \leq X \leq b) = \sum_{a \leq i \leq b} P(X = i) \]

**p.d.f.**

\[ P(a \leq X \leq b) = \int_{a}^{b} f(x) \, dx \]

**BUT NOTE:** \( f(x) \) is NOT a probability \( \forall \)  

E.g. can have \( f(x) > 1 \)

Instead, \( f(x) \) is the **probability density** at \( x \)
Probability Density

$Pr \{ x \leq X \leq x + dx \} = \int_x^{x+dx} f(x) \, dx \approx f(x) \, dx$

$f(x) = \text{"probability per unit length" at } x$ (density)
Definition: The cumulative distribution function (c.d.f.) of a continuous r.v. $X$ is defined by

$$F(x) := \Pr \{ X \leq x \} = \int_{-\infty}^{x} f(z) \, dz$$

Note:
- $F(x)$ increases monotonically to 1 as $x \to \infty$
- $f(x) = \frac{dF(x)}{dx}$
- Can use either $f(x)$ or $F(x)$ to define r.v. $X$
Example: Wheel of fortune

\[ f(x) = \begin{cases} 
0 & x < 0 \\
\frac{1}{l} & 0 \leq x \leq l \\
0 & x > l 
\end{cases} \]

\[ F(x) = \begin{cases} 
0 & x < 0 \\
\frac{x}{l} & 0 \leq x \leq l \\
1 & x > l 
\end{cases} \]
**Expectation and Variance**

**Defn:** The **expectation** of a continuous r.v. $X$ with pdf $f$ is

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

[Compare: $E[X] = \sum_a a \cdot Pr[X=a]$]

**Defn:** The **variance** of $X$ is

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

$$= \int_{-\infty}^{\infty} x^2 f(x) \, dx - E[X]^2$$

**Generally:** For a function $G: \mathbb{R} \to \mathbb{R}$,

$$E[G(X)] = \int_{-\infty}^{\infty} G(x) f(x) \, dx$$
Example: Wheel of fortune

\[
f(x) = \begin{cases} \frac{1}{l} & 0 < x \leq l \\ 0 & x < 0 \text{ or } x > l \end{cases}
\]

\[
E[X] = \int_{-\infty}^{\infty} x f(x) \, dx = \int_{0}^{l} \frac{x}{l} \, dx = \frac{x^2}{2l} \bigg|_{0}^{l} = \frac{l}{2}
\]

\[
\text{Var}(X) = \int_{-\infty}^{\infty} x^2 f(x) \, dx - E[X]^2
\]

\[
= \int_{0}^{l} \frac{x^2}{l} \, dx = \frac{x^3}{3l} \bigg|_{0}^{l} = \frac{l^2}{3}
\]

\[
\Rightarrow \text{Var}(X) = E[X^2] - E[X]^2 = \frac{l^2}{3} - \frac{l^2}{4} = \frac{l^2}{12}
\]
Compare: discrete uniform distribution on \([0, \ell-1]\) (assuming \(\ell\) integer)

i.e., \(\Pr [X=i] = \frac{1}{\ell}\) for \(i = 0, 1, \ldots, \ell-1\)

\[
E[X] = \frac{1}{\ell} \left[ 0 + 1 + 2 + \cdots + \ell - 1 \right] = \frac{1}{\ell} \cdot \frac{\ell(\ell-1)}{2} = \frac{\ell-1}{2}
\]

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

\[
E[X^2] = \frac{1}{\ell} \left[ 0 + 1^2 + 2^2 + \cdots + (\ell-1)^2 \right]
\]

\[
= \frac{1}{\ell} \cdot \frac{(\ell-1)\ell(2\ell-1)}{6} = \frac{(\ell-1)(2\ell-1)}{6}
\]

\[
\Rightarrow \text{Var}(X) = \frac{(\ell-1)(2\ell-1)}{12} - \frac{(\ell-1)^2}{4} = \frac{\ell^2-1}{12}
\]
Markov’s Inequality

**Thm:** For a continuous r.v. with p.d.f. \( f \) satisfying \( f(x) = 0 \) for \( x < 0 \):

\[
\Pr [X \geq c] \leq \frac{E[X]}{c}
\]

Chebyshev’s Inequality

**Thm:** For a continuous r.v. \( X \):

\[
\Pr [|X - E[X]| \geq c] \leq \frac{\text{Var}(X)}{c^2}
\]
Joint Distributions

**Defn:** A joint density function for two r.v.'s $X$, $Y$ is a function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ satisfying:

- $f(x,y) \geq 0 \ \forall x, y \in \mathbb{R}$
- $\iint_{-\infty}^{\infty} \iint_{-\infty}^{\infty} f(x,y) \, dx \, dy = 1$

The joint distribution of $X$, $Y$ is then

$$P \{a \leq X \leq b, \ c \leq Y \leq d\} = \int_c^d \int_a^b f(x,y) \, dx \, dy$$

**Interpretation of $f(x,y)$:**
prob. density per unit area at $(x,y)$
**Example:** Two-round game

- **Round 1:** You stake $\ell$ and win amount $X$ uniform in $[0, \ell]$
- **Round 2:** You stake $\ell X$ and win amount $Y$ uniform in $[0, X]$

- $f(x, y) = 0$ outside red triangle
- Density of $x$ is uniform on $[0, \ell]$
- Given $x$, density of $y$ is uniform on $[0, x]$

$$f(x, y) = \begin{cases} \frac{1}{\ell}x & \text{for } (x, y) \in \Delta \\ 0 & \text{otherwise} \end{cases}$$
• $f(x, y) = 0$ outside red

• Density of $x$ is uniform on $[0, l]$

• Given $x$, density of $y$ is uniform on $[0, x]$

• $f(x, y) = \begin{cases} \frac{1}{lx} & \text{for } (x, y) \in \triangle \\ 0 & \text{otherwise} \end{cases}$

Check: $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) \, dx \, dy = \int_{0}^{l} \left( \int_{0}^{x} \frac{1}{lx} \, dy \right) \, dx$

$= \int_{0}^{l} \left( \frac{y}{lx} \bigg|_{0}^{x} \right) \, dx$

$= \int_{0}^{l} \frac{1}{lx} \, dx = \frac{x}{lx} \bigg|_{0}^{l} = 1$
\[ f(x, y) = \begin{cases} \frac{1}{2}x & \text{for } (x, y) \in \Delta \\ 0 & \text{otherwise} \end{cases} \]

\[
E[Y] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} y f(x, y) \, dx \, dy = \int_{0}^{l} \int_{0}^{x} \frac{y}{t_{2x}} \, dy \, dx
\]

\[
= \int_{0}^{l} \left. \left( \frac{y^2}{2t_{2x}} \right) \right|_{0}^{x} \, dx
\]

\[
= \int_{0}^{l} \frac{x}{2l} \, dx
\]

\[
= \frac{x^2}{4l} \bigg|_{0}^{l} = \frac{l}{4}
\]
Independence

**Defn:** Continuous r.v.'s $X, Y$ are **independent** if

$$
Pr[a \leq X \leq b, c \leq Y \leq d] = Pr[a \leq X \leq b] \cdot Pr[c \leq Y \leq d]
$$

\forall a < b, c < d

**Thm:** If $X, Y$ are independent with pdf's $f(x)$, $g(y)$ respectively, then their joint density $h(x, y)$ is given by

$$
h(x, y) = f(x)g(y) \quad \forall x, y \in \mathbb{R}
$$
Application: Buffon's Needle

- Board with lines dist. \( l \) apart
- Needle length \( l \)
- Throw needle randomly onto board
- Let \( X = \begin{cases} 1 & \text{if needle hits a line} \\ 0 & \text{otherwise} \end{cases} \)

Claim: \( E[X] = \frac{2}{\pi} \)
\[ X = \begin{cases} 1 & \text{if needle hits a line} \\ 0 & \text{otherwise} \end{cases} \]

**Claim:** \( E[X] = \frac{2}{\pi} \)

If Claim is true then we can estimate \( \pi \) as in previous lecture!

Perform experiment \( N \) times \( \rightarrow X_1, \ldots, X_N \) (i.i.d.)

Output \( \hat{\rho} = \frac{1}{N} (X_1 + \ldots + X_N) \)

Then \( E[\hat{\rho}] = \frac{2}{\pi} \Rightarrow \frac{2}{\hat{\rho}} \) estimates \( \pi \)

Number of trials needed for accuracy \((1 \pm \varepsilon)\pi\) with confidence \( 1 - \delta \) is (by Chebyshev) \( \leq \frac{\pi}{2} \cdot \frac{1}{\varepsilon^2 \delta} \leq \frac{2}{\varepsilon^2 \delta} \)
Outcome of throw described by 2 random variables:

\[ Y := \text{dist. between needle midpoint & closest line} \quad 0 \leq Y \leq \frac{\pi}{2} \]

\[ \Theta := \text{angle between needle & vertical} \quad -\frac{\pi}{2} \leq \Theta \leq \frac{\pi}{2} \]
Outcome of throw described by 2 random variables:

\[ Y := \text{dist. between needle midpoint & closest line} \quad 0 \leq X \leq \frac{\pi}{2} \]

\[ \Theta := \text{angle between needle & vertical} \quad -\frac{\pi}{2} \leq Y \leq \frac{\pi}{2} \]

Joint density \( f(y, \theta) \) uniform over rectangle \([0, \frac{\pi}{2}] \times [-\frac{\pi}{2}, \frac{\pi}{2}]\)

\[ \Rightarrow f(y, \theta) = \begin{cases} \frac{2\pi}{\ell} & (y, \theta) \in \square \\ 0 & \text{otherwise} \end{cases} \]

\[ \frac{\pi \ell}{2} = \text{area of } \square \]
\[ f(y, \theta) = \begin{cases} \frac{2}{\pi} & (y, \theta) \in \square \\ 0 & \text{otherwise} \end{cases} \]

\[ X = \begin{cases} 1 & \text{if needle hits a line} \\ 0 & \text{otherwise} \end{cases} \]

Claim: \( E[X] = \frac{2}{\pi} \)

Note that \( E[X] = \Pr[E] \) where \( E \) is event “needle hits line”

Q: When does \( E \) happen?

A: When \( y \leq \frac{t}{2} \cos \theta \)
\[ f(y, \theta) = \begin{cases} \frac{2\pi}{\ell} & (y, \theta) \in \square \\ 0 & \text{otherwise} \end{cases} \]

\[ X = \begin{cases} 1 & \text{if needle hits a line} \\ 0 & \text{otherwise} \end{cases} \]

Claim: \( E[X] = \frac{2}{\pi} \)

Note that \( E[X] = \Pr[E] \) where \( E \) is event "needle hits line".

Q: When does \( E \) happen?

A: When \( Y \leq \frac{l}{2} \cos \theta \)

So \( \Pr[E] = \Pr[Y \leq \frac{l}{2} \cos \theta] = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \int_{0}^{\frac{l}{2} \cos \theta} f(y, \theta) \, dy \, d\theta \)

\[ = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \left( \frac{2y}{\pi^2} \right) \left. \frac{1}{2} \cos \theta \right|_{0}^{\frac{l}{2} \cos \theta} \, d\theta = \frac{1}{\pi} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos \theta \, d\theta = \frac{1}{\pi} \left[ \sin \theta \right]_{-\frac{\pi}{2}}^{\frac{\pi}{2}} = \frac{2}{\pi} \]