Expectation of a random variable.

X - random variable
roll a die - average value = 3.5
flip a coin - average value = 0.5 if heads = 0 and tails = 1

Definition: If X is discrete, p.f. \( f(x) = p.f. \) of X,
Then, expectation of X is \( \mathbb{E}X = \sum x f(x) \)
For a die:

<table>
<thead>
<tr>
<th></th>
<th>1/6</th>
<th>1/6</th>
<th>1/6</th>
<th>1/6</th>
<th>1/6</th>
<th>1/6</th>
</tr>
</thead>
<tbody>
<tr>
<td>f(x)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\( \mathbb{E} = 1 \times \frac{1}{6} + \ldots + 6 \times \frac{1}{6} = 3.5 \)

Another way to think about it:

Consider each \( p_i \) as a weight on a horizontal bar.
Expectation = center of gravity on the bar.

If X - continuous, \( f(x) = p.d.f. \) then \( \mathbb{E}(X) = \int x f(x) dx \)
Example: \( X \) - uniform on [0, 1], \( \mathbb{E}(X) = \int_0^1 (x \times 1) dx = 1/2 \)

Consider \( Y = r(x) \), then \( \mathbb{E}(Y) = \sum_x r(x) f(x) \) or \( \int r(x) f(x) dx \)
p.f. \( g(y) = \sum_{\{x:y=r(x)\}} f(x) \)
\( \mathbb{E}(Y) = \sum_y y g(y) = \sum_y y \sum_{\{x:y=r(x)\}} f(x) = \sum_y \sum_{\{x:r(x)=y\}} y f(x) = \sum_y \sum_{\{x:r(x)=y\}} r(x) f(x) \)
then, can drop \( y \) since no reference to \( y \):
\( \mathbb{E}(Y) = \sum_x r(x) f(x) \)

Example: \( X \) - uniform on [0, 1]
\( \mathbb{E}(X^2) = \int_0^1 X^2 \times 1 dx = 1/3 \)

\( X_1, \ldots, X_n \) - random variables with joint p.f. or p.d.f. \( f(x_1 \ldots x_n) \)
\( \mathbb{E}(r(X_1, \ldots, X_n)) = \int r(x_1, \ldots, x_n) f(x_1, \ldots, x_n) dx_1 \ldots dx_n \)

Example: Cauchy distribution
p.d.f.:

\[
f(x) = \frac{1}{\pi(1 + x^2)}
\]

Check validity of integration:

\[
\int_{-\infty}^{\infty} \frac{1}{\pi(1 + x^2)} dx = \frac{1}{\pi} \tan^{-1}(x)\big|_{-\infty}^{\infty} = 1
\]

But, the expectation is undefined:
\[ \mathbb{E}[X] = \int_{-\infty}^{\infty} |x| \frac{1}{\pi(1 + x^2)} \, dx = 2 \int_0^{\infty} \frac{x}{\pi(1 + x^2)} = \frac{1}{2\pi} \ln(1 + x^2) |_{0}^{\infty} = \infty \]

Note: Expectation of X is defined if \( \mathbb{E}[X] < \infty \)

Properties of Expectation:

1) \( \mathbb{E}(aX + b) = a\mathbb{E}(X) + b \)
   Proof: \( \mathbb{E}(aX + b) = \int (aX + b) f(x) \, dx = a \int xf(x) \, dx + b \int f(x) \, dx = a\mathbb{E}(X) + b \)

2) \( \mathbb{E}(X_1 + X_2 + ... + X_n) = \mathbb{E}X_1 + \mathbb{E}X_2 + ... + \mathbb{E}X_n \)
   Proof: \( \mathbb{E}(X_1 + X_2) = \int (x_1 + x_2) f(x_1, x_2) \, dx_1 \, dx_2 =\)
   \[ = \int x_1 f(x_1, x_2) \, dx_1 \, dx_2 + \int x_2 f(x_1, x_2) \, dx_1 \, dx_2 =\]
   \[ = \int x_1 \int f(x_1, x_2) \, dx_2 \, dx_1 + \int x_2 \int f(x_1, x_2) \, dx_1 \, dx_2 =\]
   \[ = \int x_1 f_1(x_1) \, dx_1 + \int x_2 f_2(x_2) \, dx_2 = \mathbb{E}X_1 + \mathbb{E}X_2 \]

Example: Toss a coin n times, “T” on i: \( X_i = 1 \); “H” on i: \( X_i = 0 \).

Number of tails = \( X_1 + X_2 + ... + X_n \)

\( \mathbb{E}(\text{number of tails}) = \mathbb{E}(X_1 + X_2 + ... + X_n) = \mathbb{E}X_1 + \mathbb{E}X_2 + ... + \mathbb{E}X_n \)

\( \mathbb{E}X_i = 1 \times \mathbb{P}(X_i = 1) + 0 \times \mathbb{P}(X_i = 0) = p \), probability of tails

Expectation = \( p + p + ... + p = np \)

This is natural, because you expect np of n for p probability.

\( Y = \text{Number of tails} \), \( \mathbb{P}(Y = k) = \binom{n}{k} p^k (1-p)^{n-k} \)

\( \mathbb{E}(Y) = \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k} = np \)

More difficult to see though definition, better to use sum of expectations method.

Two functions, \( h \) and \( g \), such that \( h(x) \leq g(x) \), for all \( x \in \mathbb{R} \)

Then, \( \mathbb{E}(h(X)) \leq \mathbb{E}(g(X)) \rightarrow \mathbb{E}(g(X) - h(X)) \geq 0 \)

\( \int (g(x) - h(x)) \times f(x) \, dx \geq 0 \)

You know that \( f(x) \geq 0 \), therefore \( g(x) - h(x) \) must also be \( \geq 0 \)

If \( a \leq X \leq b \rightarrow a \leq \mathbb{E}(X) \leq \mathbb{E}(b) \leq b \)

\( \mathbb{E}(I(X \in A)) = 1 \times \mathbb{P}(X \in A) + 0 \times \mathbb{P}(X \notin A), \) for \( A \) being a set on \( \mathbb{R} \)

\( Y = I(X \in A) = \{1, \text{ with probability } \mathbb{P}(X \in A); 0, \text{ with probability } \mathbb{P}(X \notin A) = 1 - \mathbb{P}(X \in A) \}

\( \mathbb{E}(I(X \in A)) = \mathbb{P}(X \in A) \}

In this case, think of the expectation as an indicator as to whether the event happens.

**Chebyshev’s Inequality**

Suppose that \( X \geq 0 \), consider \( t > 0 \), then:

\[ \mathbb{P}(X \geq t) \leq \frac{1}{t^2} \mathbb{E}(X) \]

Proof: \( \mathbb{E}(X) = \mathbb{E}(X) I(X < t) + \mathbb{E}(X) I(X \geq t) \geq \mathbb{E}(X) I(X \geq t) \geq \mathbb{E}(t) I(X \geq t) = t \mathbb{P}(X \geq t) \)

** End of Lecture 16 **