Introduction to probability

Lectures prepared by:
Elchanan Mossel
Yelena Shvets

Follows Jim Pitman’s book:
Probability
Section 3.2
Mean of a Distribution

- The **mean** $\mu$ of a probability distribution $P(x)$ over a finite set of numbers $x$ is defined as
- The mean is the average of the these numbers weighted by their probabilities:

$$\mu = \sum_x x \times P(X=x)$$
The expectation (also expected value or mean) of a random variable $X$ is the mean of the distribution of $X$.

$$E(X) = \sum_x x \times P(X=x)$$
Two Value Distributions

• If $X$ is a Bernoulli($p$) variable over \{a,b\}, then $E(X) = pa + (1-p)b$.

• If we think of $p$ and $q$ as two masses sitting over a and b then $E(X)$ would correspond to the point of balance:
A Fair Die

Let $X$ be the number rolled with a fair die.

**Question:** What is the expected value of $X$?

We can compute $E(X)$ by definition:

$$E(X) = 1 \times \frac{1}{6} + 1 \times \frac{1}{6} + 3 \times \frac{1}{6} + 4 \times \frac{1}{6} + 5 \times \frac{1}{6} + 6 \times \frac{1}{6}.$$  

$$= 3.5$$

Alternatively, we could find the point of balance on the histogram:
Binomial(10, $\frac{1}{2}$)

Question: Let $Z$ be a variable with a binomial(10, $\frac{1}{2}$) distribution. What is $E[Z]$?

By definition:

$$E(Z) = \sum_{i=0}^{10} iP(Z=i)$$
Binomial(10, $\frac{1}{2}$)

\[E(Z) = 0 + 1 \cdot 10 \left(\frac{1}{2}\right)^{10} + 2 \cdot 45 \left(\frac{1}{2}\right)^{10} + 3 \cdot 120 \left(\frac{1}{2}\right)^{10} + 4 \cdot 210 \left(\frac{1}{2}\right)^{10} + 5 \cdot 252 \left(\frac{1}{2}\right)^{10} +
\]

\[+ 6 \cdot 210 \left(\frac{1}{2}\right)^{10} + 7 \cdot 120 \left(\frac{1}{2}\right)^{10} + 8 \cdot 45 \left(\frac{1}{2}\right)^{10} + 9 \cdot 10 \left(\frac{1}{2}\right)^{10} + 10 \left(\frac{1}{2}\right)^{10}\]

\[= \left(\frac{1}{2}\right)^{10} (10 + 90 + 360 + 840 + 1260 + 1260 + 840 + 360 + 90 + 10)\]

\[= \frac{5120}{1024} = 5\]
Binomial(10, $\frac{1}{2}$)

We could also look at the histogram:
Addition Rule

• For any two random variables $X$ and $Y$ defined over the same sample space
  \[ E(X+Y) = E(X) + E(Y). \]

• Consequently, for a sequence of random variables $X_1, X_2, \ldots, X_n$,
  \[ E(X_1+X_2+\ldots+X_n) = E(X_1) + E(X_2) + \ldots + E(X_n). \]

• Therefore the mean of $\text{Bin}(10,1/2) = 5$. 
Multiplication Rule and Inequalities

- $E[aX] = \sum_x a \times P[X = x] = a \sum_x P[X = x] = a E[x]$

- If $X \geq Y$ then $E[X] \geq E[Y]$.
- This follows since $X-Y$ is non negative and
- $E[X] - E[Y] = E[X-Y] \geq 0$. 
Sum of Two Dice

• Let T be the sum of two dice. What’s E(T)?

• The “easy” way:

\[ E(T) = \sum_t tP(T=t). \]

This sum will have 11 terms.

• We could also find the center of mass of the histogram (easy to do by symmetry).
Probability distribution for $T$. 

$P(T=t)$


Sum of Two Dice

• Or we can using the addition rule:
  \[ T = X_1 + X_2, \text{ where } X_1 = 1^{st} \text{ role}, X_2 = 2^{nd}: \]

  \[
  E(T) = E(X_1) + E(X_2) = 3.5 + 3.5 = 7.\]
Indicators associate 0/1 valued random variables to events.

**Definition:** The indicator of the event $A$, $I_A$, is the random variable that takes the value 1 for outcomes in $A$ and the value 0 for outcomes in $A^c$.  

Indicators

Suppose $I_A$ is an indicator of an event $A$ with probability $p$. 

\[ I_A = 1 \quad \text{if} \quad A \quad \text{and} \quad I_A = 0 \quad \text{otherwise} \]
Then:

\[ E(I_A) = 1 \times P(A) + 0 \times P(A^c) = P(A) = p. \]
Expected Number of Events that Occur

• Suppose there are $n$ events $A_1, A_2, \ldots, A_n$.

• Let $X = I_1 + I_2 + \ldots + I_n$ where $I_i$ is the indicator of $A_i$.

• Then $X$ counts the number of events that occur.

• By the addition rule:

$$E(X) = P(A_1) + P(A_2) + \ldots + P(A_n).$$
Repeated Trials

• Let $X^i =$ indicator of success on the $i^{th}$ coin toss ($X^i = 1$ if the $i^{th}$ coin toss = H head, and $X^i = 0$ otherwise).

• The sequence $X^1, X^2, \ldots, X^n$ is a sequence of $n$ independent variables with Bernoulli($p$) distribution over \{0,1\}.

• The number of heads in $n$ coin tosses given by $S_n = X^1 + X^2 + \ldots + X^n$.

• $E(S_n) = nE(X^i) = np$

• Thus the mean of $\text{Bin}(n,p)$ RV = np.
Expected Number of Aces

• Let $Y$ be the number of aces in a poker hand.
• Then:

$Y = I_{1\text{st ace}} + I_{3\text{rd ace}} + I_{4\text{th ace}} + I_{5\text{th ace}} + I_{2\text{nd ace}}$.
• And: $E(Y) = 5 \times P(\text{ace}) = 5 \times \frac{4}{52} = 0.385$.
• Alternatively, since $Y$ has the hypergeometric distribution we can calculate:

$$E(Y) = \sum_{x=0}^{4} y \begin{pmatrix} 4 \left\{ \begin{array}{c} 48 \\ y \\ 5-y \\ 52 \\ 5 \end{array} \right. \right.$$
Suppose $X$ is an integer valued, non-negative random variable.

Let $A_i = \{X \geq i\}$ for $i=1,2,...$;

Let $I_i$ the indicator of the set $A_i$.

Then

$$X = \sum_{i} I_i.$$
Non-negative Integer Valued RV

• The equality

\[ X(\text{outcome}) = \sum_i I_i(\text{outcome}) \]

follows since if \( X(\text{outcome}) = i \), then
\( \text{outcome} \in A_1 \cap A_2 \cap \ldots \cap A_i \) but not to \( A_j \), \( j > i \).

• So
\[ (I_1 + I_2 + \ldots + I_i + I_{i+1} + \ldots)(\text{outcome}) = 1+1+\ldots+1+0+0+\ldots = i. \]
**Tail Formula for Expectation**

Let \( X \) be a non-negative integer valued RV, then:

\[
E(X) = E\left(\sum_i I_i\right) = \sum_i E(I_i) = \sum_i P(X \geq i), \quad i=1,2,3,...
\]

<table>
<thead>
<tr>
<th></th>
<th>( P(X \geq 4) )</th>
<th>( P(X=3) )</th>
<th>( P(X=4) )</th>
<th>( P(X \geq 1) )</th>
<th>( E(X) = )</th>
</tr>
</thead>
<tbody>
<tr>
<td>P(X \geq 4)</td>
<td>( P(X=4) )</td>
<td>( P(X=3) )</td>
<td>( P(X=4) )</td>
<td>( P(X \geq 4) )</td>
<td>( 4*P(X=4) )</td>
</tr>
<tr>
<td>P(X \geq 3)</td>
<td>( P(X=3) )</td>
<td>( P(X=4) )</td>
<td>( P(X \geq 3) )</td>
<td>( P(X \geq 3) )</td>
<td>( 3*P(X=3) )</td>
</tr>
<tr>
<td>P(X \geq 2)</td>
<td>( P(X=2) )</td>
<td>( P(X=3) )</td>
<td>( P(X=4) )</td>
<td>( P(X \geq 2) )</td>
<td>( 2*P(X=2) )</td>
</tr>
<tr>
<td>P(X \geq 1)</td>
<td>( P(X=1) )</td>
<td>( P(X=2) )</td>
<td>( P(X=3) )</td>
<td>( P(X \geq 1) )</td>
<td>( 1*P(X=1) )</td>
</tr>
</tbody>
</table>
Minimum of 10 Dice

Suppose we roll a die 10 times and let $X$ be the minimum of the numbers rolled.

Here $X = 2$.

**Question:** what's the expected value of $X$?
Minimum of 10 Dice

• Let’s use the tail formula to compute $E(X)$:

$$E(X) = \sum_i P(X \geq i).$$

$P(X \geq 1) = 1$;
$P(X \geq 2) = \left(\frac{5}{6}\right)^{10}$;
$P(X \geq 3) = \left(\frac{4}{6}\right)^{10}$;
$P(X \geq 4) = \left(\frac{3}{6}\right)^{10}$;
$P(X \geq 5) = \left(\frac{2}{6}\right)^{10}$;
$P(X \geq 6) = \left(\frac{1}{6}\right)^{10}$

$E(X) = \frac{6^{10} + 5^{10} + 4^{10} + 3^{10} + 4^{10} + 3^{10}}{6^{10}} = 1.17984$
Indicators

• If the events $A_1, A_2, ..., A_j$ are mutually exclusive then

$$I_1 + I_2 + ... + I_j = I_{A_1 \cup A_2 \cup ... \cup A_j}$$

• And

$$P(\bigcup_{i=1}^{j} A_i) = \sum_{i} P(A_i).$$
Tail Formula for Expectation

Let \( X \) be a non-negative integer valued RV, then:

\[
E(X) = E\left(\sum_i I_i\right) = \sum_i E(I_i)
\]

\[
E(X) = \sum_i P(X \geq i), \quad i=1,2,3,...
\]
Boole's Inequality

For a non-negative integer valued $X$ we can obtain Boole's inequality:

$$P(X \geq 1) \leq \sum_i P(X \geq i) = E(X)$$
Markov’s Inequality

Markov inequality:
If $X \geq 0$, then for every $a > 0$

$$P(X \geq a) \leq \frac{E(X)}{a}.$$ 

• This is proven as follows.
• Note that if $X \geq Y$ then $E(X) \geq E(Y)$.
• Take $Y = \text{indicator of the event } \{X \geq a\}$.
• Then $E(Y) = P(X \geq a)$ and $X \geq aY$ so:
• $E(X) \geq E(aY) = aE(Y) = aP(X \geq a)$.
Expectation of a Function of a Random Variable

• For any function $g$ defined on the range space of a random variable $X$ with a finite number of values

$$E[g(X)] = \sum_x g(x) \, P(X=x).$$

**Proof:**

• Note that:

$$P(g(X)=y) = \sum_{\{x: g(x)=y\}} P(X=x).$$

• Therefore:

$$E[g(X)] = \sum_y y \, P(g(X)=y) = \sum_y \sum_{\{x: g(x)=y\}} g(x) P(X=x) = \sum_x g(x) \, P(X=x).$$
Expectation of a Function of a Random Variable

• **Constants:**
  \[ g(X) = c \Rightarrow E[g(x)] = c. \]

• **Linear functions:**
  \[ g(X) = aX + b \Rightarrow E[g(x)] = aE(X) + b. \]

(These are the only cases when \( E(g(X)) = g(E(X)) \).)
Expectation of a Function of a Random Variable

- **Monomials:**
  \[ g(X) = X^k \Rightarrow E[g(x)] = \sum_x x^k P(X=x). \]

\[ \sum_x x^k P(X=x) \text{ is called the } k^{th} \text{ moment of } X. \]
Expectation of a Function of a Random Variable

**Question:** For $X$ representing the number on a die, what is the second moment of $X$?

$$\sum_x x^2 P(X=x) = \sum_x x^2 / 6 = \frac{1}{6} (1 + 4 + 9 + 16 + 25 + 36)$$
$$= \frac{91}{6} = 15.16667$$
Expectation of a Function of Several Random Variables

• If X and Y are two random variables we obtain:
  \[ E(g(X,Y)) = \sum_{\{\text{all } (x,y)\}} g(x,y)P(X=x, Y=y). \]

This allows to prove that \( E[X+Y] = E[X] + E[Y] \):

\[
\begin{align*}
E(X) &= \sum_{\{\text{all } (x,y)\}} x P(X=x, Y=y); \\
E(Y) &= \sum_{\{\text{all } (x,y)\}} y P(X=x, Y=y); \\
E(X+Y) &= \sum_{\{\text{all } (x,y)\}} (x+y) P(X=x, Y=y); \\
E(X+Y) &= E(X) + E(Y)
\end{align*}
\]
Expectation of a Function of Several Random Variables

\[ E(g(X,Y)) = \sum_{\text{all } (x,y)} g(x,y)P(X=x, Y=y). \]

Product:

\[ E(XY) = \sum_{\text{all } (x,y)} xy P(X=x, Y=y); \]
\[ E(XY) = \sum_x \sum_y xy P(X=x, Y=y); \]
\[ \text{Is } E(XY) = E(X)E(Y) ? \]
Product Rule for Independent Random Variables

• However, if $X$ and $Y$ are independent,
  \[ P(X=x,Y=y) = P(X=x)P(Y=y) \]
  then product formula simplifies:
  \[
  E(XY) = \sum_x \sum_y xy \ P(X=x) \ P(Y=y) \\
  = \left( \sum_x x \right) \left( \sum_y y \right) \ P(X=x) \ P(Y=y) \\
  = E(X) \ E(Y)
  \]

• If $X$ and $Y$ are independent then:
  \[
  E(XY) = E(X) \ E(Y);
  \]
Expectation interpretation as a Long-Run Average

• If we repeatedly sample from the distribution of $X$ then $P(X=x)$ will be close to the observed frequency of $x$ in the sample.

• $E(X)$ will be approximately the long-run average of the sample.
Mean, Mode and Median

• The **Mode** of \( X \) is the most likely possible value of \( X \).
• The mode need not be unique.

• The **Median** of \( X \) is a number \( m \) such that both \( P(X \leq m) \geq \frac{1}{2} \) and \( P(X \geq m) \geq \frac{1}{2} \).
• The median may also not be unique.

• **Mean** and **Median** are not necessarily possible values (mode is).
Mean, Mode and Median

For a symmetrical distribution, which has a unique Mode, all three: Mean, Mode and Median are the same.

mean = mode = median
For a distribution with a long right tail, the Mean is greater than the Median.
Roulette

Bet
- Straight Up (one number) 35:1
- Split 17:1
- Line/Street (three numbers) 11:1
- Corner (four numbers) 8:1
- Five line (5 numbers-0,00,1,2 or 3) 6:1
- Six line (six numbers) 5:1
- Column (twelve numbers) 2:1
- Dozens (twelve numbers) 2:1
- 1 - 18 1:1
- Even 1:1
- Red 1:1
- Black 1:1
- Odd 1:1
- 19-36 1:1
Betting on Red

Suppose we want to be $1 on Red. Our chance of winning is $18/38$.

Question:
What should be the pay-off to make it a fair bet?
Betting on Red

This question really only makes sense if we repeatedly bet $1 on Red.

Suppose that we could win $x if Reds come up and lose $1, otherwise. If $X$ denotes our returns then $P(X=x) = 18/38; P(X=-1)=20/38$.

In a fair game, we expect to break even on average.
Betting on Red

Our expected return is:

\[ x \times \frac{18}{38} - 1 \times \frac{20}{38}. \]

Setting this to zero gives us

\[ x = \frac{20}{18} = 1.1111111\ldots. \]

This is greater than the pay-off of 1:1 that is offered by the casino.