Outline

Summary of Previous Lecture
Probability Definition
Properties of Probability
Random Variables
Probability Distributions

Expectation
Mean and Variance
Conditional Probability and Bayes’ Rule
Part I

Review of Previous Lecture
Elements of Probability

- Fixed outcome space.
- Independent trials.
- Identical trials (all have same chance of “success”).
- Infinite number of trials.
Rules of Probability

- $0 \leq \Pr(A) \leq 1$.
- $\Pr(S) = 1$.
- $\Pr(A^c) = 1 - \Pr(A)$.
- If $A$ and $B$ are disjoint, then $\Pr(A \text{ or } B) = \Pr(A) + \Pr(B)$.
- If $A$ and $B$ are independent, then $\Pr(A \text{ and } B) = \Pr(A) \Pr(B)$.

**Tip:** To calculate probability for a complex event: decompose it into simpler (disjoint and/or independent) events, and then apply the rules.
Random Variables

- Classification:
  - Qualitative: binary, nominal, ordinal.
  - Quantitative: discrete, continuous.

- Define variables carefully.
For integer-valued (qualitative and discrete) variables, the \textit{probability mass function} (pmf) gives the probability of each value.

For real-valued (continuous) variables, the \textit{probability density function} (pdf) is used such that the area under curve gives the probability of falling into any particular interval.
Part II

Expectation
For $n$ samples, $x_1, x_2, \ldots, x_n$, the *sample* mean and variance are defined as:

\[
\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i,
\]

\[
s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2.
\]
Sample Mean and Variance

Example

This is a sample from tossing a die:

\[
\begin{array}{cccccccc}
1 & 1 & 1 & 3 & 6 & 1 & 5 & 1 & 6 & 6 \\
2 & 2 & 3 & 2 & 4 & 3 & 1 & 2 & 2 & 6 \\
\end{array}
\]

Compute the mean and variance.
Mean and Frequency Table

<table>
<thead>
<tr>
<th>value</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>counts</td>
<td>6</td>
<td>5</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td>frequency</td>
<td>0.3</td>
<td>0.25</td>
<td>0.15</td>
<td>0.05</td>
<td>0.05</td>
<td>0.2</td>
</tr>
</tbody>
</table>

\[
\bar{x} = \frac{1}{20} (6 \times 1 + 5 \times 2 + 3 \times 3 + 1 \times 4 + 1 \times 5 + 4 \times 6)
\]

\[
= 0.3 \times 1 + 0.25 \times 2 + 0.15 \times 3 + 0.05 \times 4 + 0.05 \times 4 + 0.2 \times 6
\]

\[
= 2.9
\]
Computing the Sample Mean for Discrete Data

If $X$ can take $k$ values, then

$$
\bar{x} = \sum_{i=1}^{k} x_i \times f_i,
$$

where $f_i$ is the frequency of $x_i$. 

PubH 6450 Fall 2006
Expectation

Definitions

- If a discrete random variable $X$ has pmf $p(x) = \Pr(X = x)$, then the *expectation* of $X$ is defined as:

$$E[X] = \sum_x x \cdot p(x).$$

- In general, the expectation of $f(X)$ (any function of a random variable is also a random variable) is:

$$E[f(X)] = \sum_x f(x)p(x).$$
Population (Theoretical) Moments and Variance

- \( \mathbb{E}[X^2] = \sum x^2 p(x) \).
- The expectation of \( X^n \) is called the \( n \)th moment of \( X \).
- The first moment of \( X \) is often called the population mean, and denoted by the Greek letter \( \mu \).
- The \( n \)th central moment is the expectation of \( (X - \mathbb{E}X)^n \).
- The population variance is the second central moment,

\[
\text{Var}[X] = \mathbb{E}(X - \mathbb{E}X)^2 = \sum_x (x - \mathbb{E}X)^2 p(x),
\]

and is often denoted by the Greek letter \( \sigma^2 \).
Let $\mu = E X$, and note that

$$\sigma^2 = \text{Var}[X] = \sum_x (x - \mu)^2 p(x)$$

$$= \sum_x (x^2 - 2x\mu - \mu^2) p(x)$$

$$= \sum_x x^2 p(x) - 2\mu \sum_x xp(x) + \sum_x \mu^2 p(x)$$

$$= E X^2 - 2\mu \sum_x xp(x) + \mu^2 \sum_x p(x)$$

$$= E X^2 - \mu^2.$$
Rules of Means and Variances

Derive the following rules:

- \( E[aX] = a \, E[X] \).
- \( E[aX + bY] = a \, E[X] + b \, E[Y] \).
- \( \text{Var}[aX] = a^2 \, \text{Var}[X] \).
- If \( X \) and \( Y \) are independent, then
  \[ \text{Var}[X + Y] = \text{Var}[X] + \text{Var}[Y] \.
  \]

Note: To get this last rule, we need to use the fact that if \( X \) and \( Y \) are independent, then

\[ E(XY) = E(X)E(Y) \, . \]

(Note similarity to Probability Rule 5.)
Mean and Variance of an Average

If $X$ and $Y$ are independent and identically distributed (i.i.d), with mean $\mu$ and variance $\sigma^2$, let $Z = (X + Y)/2$, what is the mean and variance of $Z$?

**Answer:** Use rules on the previous page to get:

\[
\begin{align*}
E Z &= \mu \\
\text{Var } Z &= \frac{1}{2} \sigma^2
\end{align*}
\]
Mean and Variance for a Die or Coin Toss

Examples

- Compute the mean and variance of a fair die.
- Recall that the pmf for Bernoulli random variable is

\[ \Pr(X = x) = p^x(1 - p)^{1-x}, \quad x = 0, 1. \]

Compute the mean and variance of \( X \).

Answer here turns out to be \( E(X) = p \) and 
\( Var(X) = p(1 - p) \)...
Plot of $\text{Var}(X) = p(1 - p)$ for $X \sim \text{Bernoulli}$:

Uncertainty is highest when $p = 0.5$, which makes sense!
Noncubic Die Revisited

- Which values of $r$ to use?
- How many times to roll each die? Should it depend on $r$?
Definitions

- A *statistic* is a function of the data, e.g., sample mean ($\bar{x}$), sample variance ($s^2$). Statistics are random variables.

- For a random variable $X$, the expectation of $f(X)$, e.g., mean ($\mu$) and variance ($\sigma^2$), are *properties* (parameters) of the distribution.
Bernoulli statistics

- Since statistics are random variables, they have corresponding distributions.
- Consider an experiment of tossing a coin $n$ times, and getting $X$ heads. Let $\hat{p} = \frac{X}{n}$, the proportion of heads. Note that this is the sample mean of the Bernoulli (success/failure) random variable.
- What is the distribution of $\hat{p}$? How does it change with $n$?
A more general multiplication rule

If two events $A$ and $B$ are not independent, we can still compute the probability both occur as

$$P(A \text{ and } B) = P(A)P(B|A),$$

where $P(B|A)$ denotes the conditional probability that $B$ occurs given that $A$ has already occurred.

Example

What is the chance of drawing 2 aces from a full deck of cards? Let $A = \text{ace on draw 1}$, and $B = \text{ace on draw 2}$. Then

$$P(\text{two aces}) = P(\text{ace on draw 1})P(\text{ace on draw 2}|\text{ace on draw 1})$$

$$= \left(\frac{4}{52}\right)\left(\frac{3}{51}\right) = \frac{12}{2652} \approx .0045$$
Notice that if we need to compute a conditional probability and already know the probability that both events occur, we can flip the previous formula around to get:

**Definition**

The conditional probability of an event $B$ given another event $A$ is

$$P(B|A) = \frac{P(A \text{ and } B)}{P(A)},$$

provided $P(A) > 0$. 
Sometimes we need to compute a conditional probability that conditions the events \textit{in the opposite order}. In this case, a useful result may be \textit{Bayes’ Rule}:

\textbf{Definition}

If $A$ and $B$ are any events whose probabilities are not 0 or 1, then

$$P(A|B) = \frac{P(B|A)P(A)}{P(B|A)P(A) + P(B|A^c)P(A^c)}$$

Named for the Rev. Thomas Bayes, a 18th century English nonconformist minister and amateur mathematician!
Example of Bayes’ Rule

Suppose we have a diagnostic test for a disease $D$ whose prevalence in the population is 10%. When given to a diseased individual, the test is positive with probability .95; when given to a healthy individual, the test is positive with probability .15 (false positive). Find the probability that a randomly selected individual with a positive test really has the disease.

Solution: Use Bayes’ Rule with

$$A = \text{has disease } D$$

and

$$B = \text{test is +}$$
Example of Bayes’ Rule

Applying Bayes’ Rule, we get:

\[
P(D|\text{test}+) = \frac{P(\text{test} + | D)P(D)}{P(\text{test} + | D)P(D) + P(\text{test} + | \text{no } D)P(\text{no } D)}
\]

\[
= \frac{(.95)(.10)}{(.95)(.10) + (.15)(.90)}
\]

\[
= \frac{.095}{.095 + .135} = .413
\]

\[\rightarrow\] Even though the test is pretty good, the rarity of the disease in the population means a positive test still implies a less than 50% chance of actually having the disease!