Expectation

Suppose a discrete random variable with following probability distribution:

<table>
<thead>
<tr>
<th>x</th>
<th>100</th>
<th>200</th>
</tr>
</thead>
<tbody>
<tr>
<td>P</td>
<td>0.01</td>
<td>0.99</td>
</tr>
</tbody>
</table>

• What is the ”average value” of $X$ ?
Expectation (Cont’d)

• The expected value of $h(X)$, a function of discrete random variable $X$ is: $E(h(X)) = \sum_x h(x)P(x)$
  
  Example 1: $Var(X) = E(X - \mu)^2$
  
  Example 2: $X$ is Bernoulli variable. $E(X^2) = p$

• $E(.)$ is a linear operator: $E(c_0 + \sum_{i=1}^{n} c_i X_i) = c_0 + \sum_{i=1}^{n} c_i E(X_i)$
  
  Example 1: $Y \sim B(n, p); Y = X_1 + \cdots + X_n$, where $X_i$ are Bernoulli variable. $E(X_i) = p$, hence $E(Y) = E(X_1) + \cdots + E(X_n) = np$
  
  Example 2: $X_1, \cdots X_n$ are a simple random sample (i.e., independent and identically distributed, or iid), $E(X_i) = \mu, \forall i$, $E(X) = \mu$, therefore $\bar{X}$ is an unbiased estimator of $\mu$. 
Expectation (Cont’d)

• $Cov(X, Y) \equiv \sigma_{xy} = E(X - \mu_x)(Y - \mu_y)$ Note: $Cov(X, X) = Var(X)$

• Definition: $X_1, \ldots, X_n$ are independent random variables, if $P(X_1 \leq x_1, \ldots, X_n \leq x_n) = P(X_1 \leq x_1) \cdots P(X_n \leq x_n)$ for any $x_1, \ldots, x_n$

If $X, Y$ are independent, $E(XY) = E(X)E(Y)$

• If $X, Y$ are independent, then $Cov(X, Y) = 0$

Proof: $Cov(X, Y) = E(X - \mu_x)(Y - \mu_y) = E(XY) - \mu_x \mu_y = E(X)E(Y) - \mu_x \mu_y = 0$
**Variance**

- \( \text{Var}(X) = E(X^2) - E(X)^2 \)
  
  Proof: \( \text{Var}(X) = E(X^2 - 2\mu X + \mu^2) = E(X^2) - 2\mu E(X) - \mu^2 = E(X^2) - \mu^2 \)
  
  For Bernoulli variable \( X \), \( \text{Var}(X) = E(X^2) - E(X)^2 = p - p^2 = p(1 - p) \)

- \( \text{Var}(a + bX) = b^2 \text{Var}(X) \)

- \( \text{Var}(\sum_{i=1}^{n} c_i X_i) = \sum_{i=1}^{n} c_i c_j \text{Cov}(X_i, X_j) \)

- If \( X_1, \cdots, X_n \) are independent random variables,
  
  \[ \text{Var}\left(\sum_{i=1}^{n} c_i X_i\right) = \sum_{i=1}^{n} c_i^2 \text{Var}(X_i) \]

  Example: \( Y \sim B(n, p) \); \( Y = X_1 + \cdots + X_n \); \( \text{Var}(Y) = np(1-p) \)
Binomial

You are taking a multiple-choice quiz that consists of five questions. Each question has four possible answers, only one of which is correct. To complete the quiz, you randomly guess the answer to each question. Find the probability of guessing

• exactly three answers correctly.
• at least three answers correctly.
• less than three answers correctly.
Figure 1: Binomial Probability Function for different $p$
Skewness

• A distribution can be symmetric or skewed. For symmetric distribution, mean = median.

• A distribution is right-skewed if its tail extends to the right. In this case, mean > median.

• A distribution is left-skewed if its tail extends to the left. In this case, mean < median.
Binomial (cont’d)

**Application:** Suppose you work for a marketing agency and are in charge of creating a television ad for Brand A toothpaste. The toothpaste manufacturer claims that 40% of toothpaste buyers prefer its brand. To check whether the manufacturer’s claim is reasonable, your agency conducts a survey. Of 100 toothpaste buyers selected at random, you find that only 35 (or 35%) prefer Brand A. Could the manufacturer’s claim still be true? What if only 25 people (or 25%) prefer Brand A?
Binomial Probability Function with $P=0.40$, $n=100$
Poisson

The mean number of accidents per month at a certain intersection is 3.

- What is the probability that in any given month 4 accidents will occur at this intersection?
- What is the probability that more than four accidents will occur in any given month at the intersection?
Figure 2: Poisson Probability Function for different $\mu$. 
Poisson as a proximation to Binomial

The Poisson distribution is a good approximation for the Binomial when number of trial, \( n \) is large (say, \( n \geq 20 \)) and probability of success \( p \) is small (say, \( p \leq 0.05 \)). In this case, \( \mu = np \)

**Example:** A company is opening a large regional office, and each of its 200 managers is allowed to order his or her own choice of a telephone. Assuming independence of choices, and that each of the 1,000 different combinations of telephone are equally likely, what is the probability that a particular choice will be made by none, one, two, or three of the managers?

Solution: \( X \sim B(200, 0.001) \), Since \( n = 200 \) is large and \( p = 1/1000 \) is small, the Poisson approxmiation should work well. The parameter is \( \mu = np = 0.2 \). We get \( P(X = 0) \approx e^{-0.2} = 0.8187 \);
Hypergeometric

Hypergeometric distribution is the correct model for sampling without replacement. However, when $n/N \leq 5\%$, Hypergeometric variable can be approximated well by a Binomial variable with $p = M/N$. 
Binomial vs Hypergeometric $p = 0.25$, $n = 4$, $n/N = 0.2$