Lecture 12

- **Readings:** Section 4.3; parts of Section 4.5 (mean and variance only; no transforms)

Lecture outline

- Conditional expectation
  - Law of iterated expectations
  - Law of total variance
- Sum of a random number of independent r.v.'s
  - mean, variance

Conditional expectations

- Given the value $y$ of a r.v. $Y$:
  \[ E[X | Y = y] = \sum_x x p_{X|Y}(x | y) \]
  (integral in continuous case)
- Stick example: stick of length $\ell$ break at uniformly chosen point $Y$ break again at uniformly chosen point $X$
  \[ E[X | Y = y] = \frac{y}{2} \] (number)
  \[ E[X | Y] = \frac{Y}{2} \] (r.v.)
- **Law of iterated expectations:**
  \[ E[E[X | Y]] = \sum_y E[X | Y = y] p_Y(y) = E[X] \]
- In stick example:
  \[ E[X] = E[E[X | Y]] = E[Y/2] = \ell/4 \]

var($X \mid Y$) and its expectation

- var($X \mid Y = y$) = $E[(X - E[X \mid Y = y])^2 | Y = y]$
- var($X \mid Y$): a r.v.
  - with value var($X \mid Y = y$) when $Y = y$
- **Law of total variance:**
  \[ \text{var}(X) = E[\text{var}(X \mid Y)] + \text{var}(E[X \mid Y]) \]

Proof:

(a) Recall: \( var(X) = E[X^2] - (E[X])^2 \)

(b) \( var(X \mid Y) = E[X^2 \mid Y] - (E[X \mid Y])^2 \)

(c) \( E[\text{var}(X \mid Y)] = E[X^2] - E[(E[X \mid Y])^2] \)

(d) \( \text{var}(E[X \mid Y]) = E[(E[X \mid Y])^2] - (E[X])^2 \)

Sum of right-hand sides of (c), (d):
\[ E[X^2] - (E[X])^2 = \text{var}(X) \]

Section means and variances

Two sections:
\[ y = 1 \) (10 students); \( y = 2 \) (20 students)

\[ y = 1 : \frac{1}{10} \sum_{i=1}^{10} x_i = 90 \]
\[ y = 2 : \frac{1}{20} \sum_{i=11}^{30} x_i = 60 \]

\[ E[X] = \frac{1}{30} \sum_{i=1}^{30} x_i = \frac{90 \cdot 10 + 60 \cdot 20}{30} = 70 \]
\[ E[X \mid Y = 1] = 90, \quad E[X \mid Y = 2] = 60 \]

\[ E[X \mid Y] = \begin{cases} 90, & \text{w.p. } 1/3 \\ 60, & \text{w.p. } 2/3 \end{cases} \]
\[ E[E[X \mid Y]] = \frac{1}{3} \cdot 90 + \frac{2}{3} \cdot 60 = 70 = E[X] \]

\[ \text{var}(E[X \mid Y]) = \frac{1}{3}(90 - 70)^2 + \frac{2}{3}(60 - 70)^2 \]
\[ = \frac{600}{3} = 200 \]
Section means and variances (ctd.)

\[
\begin{align*}
\frac{1}{10} \sum_{i=1}^{10} (x_i - 90)^2 &= 10 \\
\frac{1}{20} \sum_{i=11}^{30} (x_i - 60)^2 &= 20 \\
\end{align*}
\]

\[
\text{var}(X | Y = 1) = 10 \quad \text{var}(X | Y = 2) = 20
\]

\[
\text{var}(X | Y) = \begin{cases} 
10, & \text{w.p. } 1/3 \\
20, & \text{w.p. } 2/3 
\end{cases}
\]

\[
\text{E}[\text{var}(X | Y)] = \frac{1}{3} \cdot 10 + \frac{2}{3} \cdot 20 = \frac{50}{3}
\]

\[
\text{var}(X) = \text{E}[\text{var}(X | Y)] + \text{var}(\text{E}[X | Y])
\]

\[
= \frac{50}{3} + 200
\]

\[
= \text{(average variability within sections)}
\]

\[
+ \text{(variability between sections)}
\]

Example

\[
\text{var}(X) = \text{E}[\text{var}(X | Y)] + \text{var}(\text{E}[X | Y])
\]

\[
\text{E}[X | Y = 1] = \quad \text{E}[X | Y = 2] = 
\]

\[
\text{var}(X | Y = 1) = \quad \text{var}(X | Y = 2) = 
\]

\[
\text{E}[X] = 
\]

\[
\text{var}(\text{E}[X | Y]) = 
\]

Sum of a random number of independent r.v.'s

- \(N\): number of stores visited
  (\(N\) is a nonnegative integer r.v.)
- \(X_i\): money spent in store \(i\)
  - \(X_i\) assumed i.i.d.
  - independent of \(N\)
- Let \(Y = X_1 + \cdots + X_N\)
  \[
  \text{E}[Y | N = n] = \text{E}[X_1 + X_2 + \cdots + X_n | N = n] \\
  = \text{E}[X_1 + X_2 + \cdots + X_n] \\
  = \text{E}[X_1] + \text{E}[X_2] + \cdots + \text{E}[X_n] \\
  = n \text{E}[X]
  \]
- \(\text{E}[Y | N] = N \text{E}[X]\)

\[
\text{E}[Y] = \text{E}[\text{E}[Y | N]] = \text{E}[N \text{E}[X]] = \text{E}[N] \text{E}[X]
\]

Variance of sum of a random number of independent r.v.'s

- \(\text{var}(Y) = \text{E}[\text{var}(Y | N)] + \text{var}(\text{E}[Y | N])\)
- \(\text{E}[Y | N] = N \text{E}[X]\)
  \[
  \text{var}(\text{E}[Y | N]) = (\text{E}[X])^2 \text{var}(N)
  \]
- \(\text{var}(Y | N = n) = n \text{var}(X)\)
  \[
  \text{var}(Y | N) = N \text{var}(X)
  \]
  \[
  \text{E}[\text{var}(Y | N)] = \text{E}[N] \text{var}(X)
  \]

\[
\text{var}(Y) = \text{E}[\text{var}(Y | N)] + \text{var}(\text{E}[Y | N])
\]

\[
= \text{E}[N] \text{var}(X) + (\text{E}[X])^2 \text{var}(N)
\]