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. $Y$:
  \[ E[X \mid Y = y] = \sum_x x p_{X \mid Y}(x \mid y) \]
  (integral in continuous case)
• Stick example: stick of length $\ell$
  break at uniformly chosen point $Y$
  break again at uniformly chosen point $X$
• $E[X \mid Y = y] = \frac{y}{2}$ (number)

\[ E[X \mid Y] = \frac{Y}{2} \quad (r.v.) \]

• Law of iterated expectations:
  \[ E[E[X \mid Y]] = \sum_y E[X \mid Y = y] p_Y(y) = E[X] \]
• In stick example:
  \[ E[X] = E[E[X \mid Y]] = E[Y/2] = \ell/4 \]

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

• $\text{var}(X \mid Y = y) = E[(X - E[X \mid Y = y])^2 \mid Y = y]$
• $\text{var}(X \mid Y)$: a r.v.
  with value $\text{var}(X \mid Y = y)$ 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: $\text{var}(X) = E[X^2] - (E[X])^2$
(b) $\text{var}(X \mid Y) = E[X^2 \mid Y] - (E[X \mid Y])^2$
(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 \quad 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.)

\[
\frac{1}{10} \sum_{i=1}^{10} (x_i - 90)^2 = 10 \quad \frac{1}{20} \sum_{i=11}^{30} (x_i - 60)^2 = 20
\]

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

\[
\begin{align*}
\text{var}(X \mid Y) &= \begin{cases} 
10, & \text{w.p. } 1/3 \\
20, & \text{w.p. } 2/3 
\end{cases} \\
E[\text{var}(X \mid Y)] &= \frac{1}{3} \cdot 10 + \frac{2}{3} \cdot 20 = \frac{50}{3}
\end{align*}
\]

\[
\text{var}(X) = E[\text{var}(X \mid Y)] + \text{var}(E[X \mid Y]) \\
= \frac{50}{3} + 200 \\
= \text{(average variability within sections)} + \text{(variability between sections)}
\]

Example

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

\[
\begin{align*}
E[X \mid Y = 1] &= \quad E[X \mid Y = 2] = \\
\text{var}(X \mid Y = 1) &= \quad \text{var}(X \mid Y = 2) = \\
E[X] &= \\
\text{var}(E[X \mid Y]) &= 
\end{align*}
\]

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 \(Y = X_1 + \cdots + X_N\)
  \[
  E[Y \mid N = n] = E[X_1 + X_2 + \cdots + X_n \mid N = n] \\
  = E[X_1 + X_2 + \cdots + X_n] \\
  = E[X_1] + E[X_2] + \cdots + E[X_n] \\
  = n E[X]
  \]

- \(E[Y \mid N] = N E[X]\)

\[
E[Y] = E[E[Y \mid N]] \\
= E[N E[X]] \\
= E[N] E[X]
\]

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

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

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

- \(\text{var}(Y \mid N = n) = n \text{var}(X)\)
  \[
  \text{var}(Y \mid N) = N \text{var}(X) \\
  E[\text{var}(Y \mid N)] = E[N] \text{var}(X)
  \]

\[
\begin{align*}
\text{var}(Y) &= E[\text{var}(Y \mid N)] + \text{var}(E[Y \mid N]) \\
&= E[N] \text{var}(X) + (E[X])^2 \text{var}(N)
\end{align*}
\]