STAT 516

Using conditioning to evaluate mean and variance

Prof. Michael Levine

April 14, 2020

Motivation

- The following approach often helps to find an expectation with respect to a complicated joint pmf:
  1. Condition X on the value y of a suitable random variable Y
  2. Compute the conditional expectation $E(X|y)$
  3. Average the resulting conditional expectation over y
- The choice of Y is crucial

The law of iterated expectations

- Let X and Y be random variables on the same probability space Ω.
- Suppose $E(X)$ and $E(X|Y=y)$ exist for each y.
- Then,
$$E(X) = E_Y[E(X|Y=y)]$$
- In the discrete case, this amounts to
$$E(X) = \sum_y \mu_X(y) p_Y(y)$$

Proof

- The discrete case is very simple: by definition,

$$\mu_X(y) = \frac{\sum_x x p(x,y)}{p_Y(y)}$$

- Thus,

$$\sum_y \mu_X(y) p_Y(y) = \sum_y \sum_x x p(x,y) = \sum_x \sum_y x p(x,y)$$
$$= \sum_x x \sum_y p(x,y) = \sum_x x p_X(x) = E(X)$$

Iterated Variance formula

- Let X and Y be random variables defined on the same probability space Ω
- Suppose $Var(X)$ and $Var(X|Y = y)$ exist for each y.
- Then,

$$Var(X) = E_Y[Var(X|Y = y)] + Var_Y[E(X|Y = y)]$$

- This formula is valid for all types of variables (not just continuous ones)

Some simple implications of the iterated variance formula

- $$Var(g(X)|X = x) = 0$$
- $$Var(g(X)h(Y)|Y = y) = h^2(y)Var(g(X)|Y = y)$$

Example I: a two-stage experiment

- Suppose n fair dice are rolled. Those that show a six are rolled again
- What are the mean and the variance of the number of sixes obtained in the second round of this experiment?
- Suppose Y is the number of dice in the first round that show a six
- Let X be the number of dice in the second round that show a six

Example I: a two-stage experiment

- Given $Y = y$, $X \sim Bin\left(y, \frac{1}{6}\right)$
- Also, $Y \sim Bin\left(n, \frac{1}{6}\right)$
- Therefore,

$$E(X) = E[E(X|Y = y)] = E_Y\left[\frac{y}{6}\right] = \frac{n}{36}$$

Example I: a two-stage experiment

- Moreover,

$$
\begin{aligned}
Var(X) &= E_Y[Var(X|Y=y)] + Var_Y[E(X|Y=y)] \\
&= E_Y\left[y\frac{1}{6}\frac{5}{6}\right] + Var_Y\left[\frac{y}{6}\right] \\
&= \frac{5}{36}\frac{n}{6} + \frac{1}{36}n\frac{1}{6}\frac{5}{6} \\
&= \frac{35n}{1296}
\end{aligned}
$$