LECTURE 7: Conditioning on a random variable; Independence of r.v.'s

- Conditional PMFs

  - Conditional expectations

  - Total expectation theorem

- Independence of r.v.'s

  - Expectation properties

  - Variance properties

- The variance of the binomial

- The hat problem: mean and variance

$$A = \{Y = y\}$$

$$p_{X|A}(x \mid A) = \mathbf{P}(X = x \mid A) \qquad p_{X|Y}(x \mid y) = \mathbf{P}(X = x \mid Y = y) = \frac{\mathbf{P}(X = x, Y = y)}{\mathbf{P}(Y = y)}$$

$$p_{X|Y}(x \mid y) = \frac{p_{X,Y}(x, y)}{p_Y(y)} \qquad \text{defined for } y \text{ such that } p_Y(y) > 0 \qquad \sum_x p_{X|Y}(x \mid y) = 1$$

$$Y = 2 \qquad P_Y(2) = 5/20$$

$$P_{X|Y}(1 \mid 2) = 0$$

$$P_{X|Y}(2 \mid 2) = 1/5$$

$$P_{X|Y}(\cdot \mid 2)$$
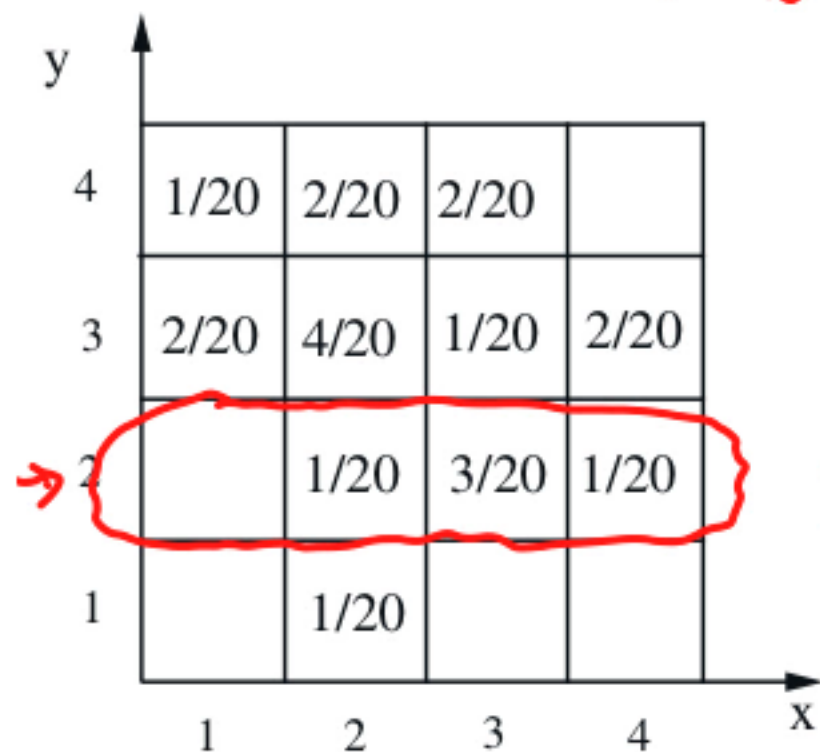
$$p_{X,Y}(x, y) = p_Y(y) \, p_{X|Y}(x \mid y)$$

$$p_{X,Y}(x, y) = p_X(x) \, p_{Y|X}(y \mid x)$$

Conditional PMFs involving more than two r.v.'s

- Self-explanatory notation

$$p_{X|Y,Z}(x \mid y, z) = P(X = x \mid Y = y, Z = z) = \frac{P(X = x, Y = y, Z = z)}{P(Y = y, Z = z)} = \frac{p_{X,Y,Z}(x, y, z)}{p_{Y,Z}(y, z)}$$

$$p_{X,Y|Z}(x, y \mid z) = P(X = x, Y = y \mid Z = z)$$

- Multiplication rule

$$\mathbf{P}(A \cap B \cap C) = \mathbf{P}(A)\,\mathbf{P}(B \mid A)\,\mathbf{P}(C \mid A \cap B)$$

$$A = \{X = x\} \qquad B = \{Y = y\} \qquad C = \{Z = z\}$$

$$p_{X,Y,Z}(x, y, z) = p_X(x)\, p_{Y|X}(y \mid x)\, p_{Z|X,Y}(z \mid x, y)$$

Conditional expectation

$$A = \{Y = y\}$$

$$\mathbf{E}[X] = \sum_x x p_X(x) \qquad \mathbf{E}[X \mid A] = \sum_x x\, p_{X|A}(x) \qquad \mathbf{E}[X \mid Y = y] = \sum_x x p_{X|Y}(x \mid y)$$

- Expected value rule

$$\mathbf{E}[g(X)] = \sum_x g(x)\, p_X(x) \qquad \mathbf{E}[g(X) \mid A] = \sum_x g(x)\, p_{X|A}(x)$$

$$\mathbf{E}[g(X) \mid Y = y] = \sum_x g(x)\, p_{X|Y}(x \mid y)$$

Total probability and expectation theorems

- $A_1, \ldots, A_n$: partition of Ω

$$Y = \{ \gamma_1, \ldots, \gamma_n \} \qquad A_i = \{ Y = \gamma_i \}$$

- $p_X(x) = \mathbf{P}(A_1)\, p_{X|A_1}(x) + \cdots + \mathbf{P}(A_n)\, p_{X|A_n}(x)$

$$p_X(x) = \sum_y p_Y(y)\, p_{X|Y}(x \mid y)$$

- $\mathbf{E}[X] = \mathbf{P}(A_1)\, \mathbf{E}[X \mid A_1] + \cdots + \mathbf{P}(A_n)\, \mathbf{E}[X \mid A_n]$

$$\boxed{\mathbf{E}[X] = \sum_y p_Y(y)\, \mathbf{E}[X \mid Y = y]} \quad \bullet$$

- Fine print:
  Also valid when Y is a discrete r.v. that ranges over an infinite set,
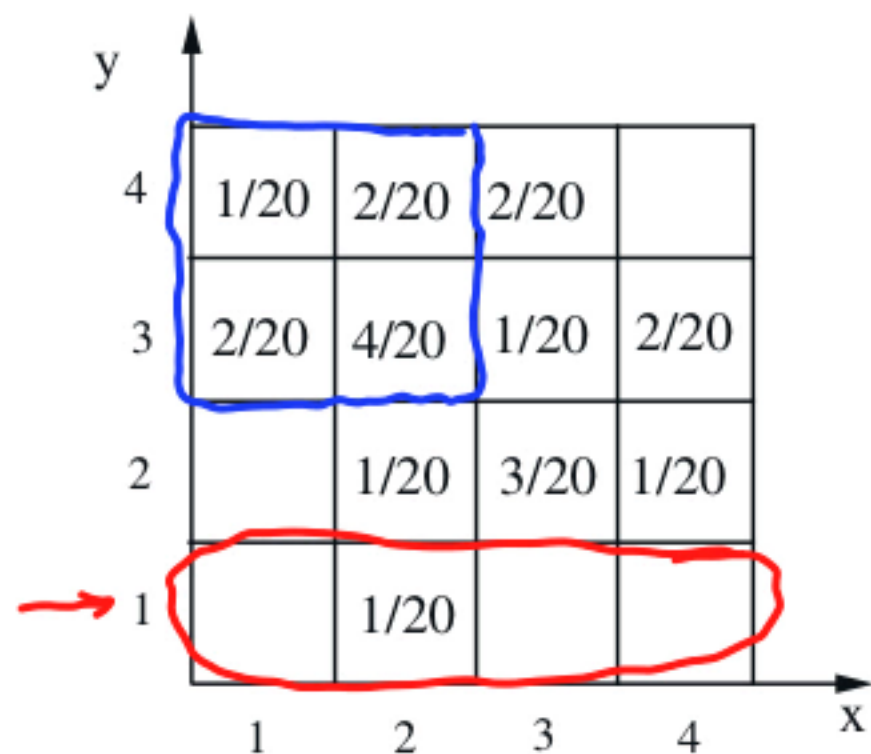 as long as $\mathbf{E}[|X|] < \infty$

Independence

- of two events: $\mathbf{P}(A \cap B) = \mathbf{P}(A) \cdot \mathbf{P}(B)$ $\mathbf{P}(A \mid B) = \mathbf{P}(A)$

- of a r.v. and an event: $\mathbf{P}(X = x \text{ and } A) = \mathbf{P}(X = x) \cdot \mathbf{P}(A),$ for all x

$$p_{X|A}(x) = p_X(x), \text{ for all } x \qquad P(A \mid X = x) = P(A), \text{ for all } x$$

- of two r.v.'s: $\mathbf{P}(X = x \text{ and } Y = y) = \mathbf{P}(X = x) \cdot \mathbf{P}(Y = y),$ for all x, y

$$p_{X|Y}(x|y) = p_X(x) \qquad p_{X,Y}(x, y) = p_X(x)\, p_Y(y), \quad \text{for all } x, y$$

$$p_{Y|X}(y|x) = p_Y(y)$$

> X, Y, Z are **independent** if:
> $p_{X,Y,Z}(x, y, z) = p_X(x)\, p_Y(y)\, p_Z(z),$ for all x, y, z

Example: independence and conditional independence

- Independent? No

$$P_X(1) = 3/20$$

$$P_{X|Y}(1 \mid 1) = 0$$

- What if we condition on $X \leq 2$ and $Y \geq 3$?

Yes

Independence and expectations

- In general: $\mathbf{E}[g(X,Y)] \neq g\big(\mathbf{E}[X], \mathbf{E}[Y]\big)$

always true

- Exceptions: $\mathbf{E}[aX + b] = a\mathbf{E}[X] + b$ $\mathbf{E}[X + Y + Z] = \mathbf{E}[X] + \mathbf{E}[Y] + \mathbf{E}[Z]$

If X, Y are **independent**: $\mathbf{E}[XY] = \mathbf{E}[X]\mathbf{E}[Y]$

$g(X)$ and $h(Y)$ are also independent: $\mathbf{E}[g(X)h(Y)] = \mathbf{E}[g(X)] \cdot \mathbf{E}[h(Y)]$

$E[g(X,Y)]$ $g(x,y) = xy$

$$= \sum_x \sum_y xy \, P_{X,Y}(x,y) = \sum_x \sum_y x \, y \, P_X(x) P_Y(y)$$

$$= \sum_x x \, P_X(x) \sum_y y \, P_Y(y) = E[X] \, E[Y.]$$

Independence and variances

- Always true: $\text{var}(aX) = a^2\text{var}(X)$ $\text{var}(X + a) = \text{var}(X)$

- In general: $\text{var}(X + Y) \neq \text{var}(X) + \text{var}(Y)$

If X, Y are independent: $\text{var}(X + Y) = \text{var}(X) + \text{var}(Y)$

assume
$$E[x] = E[Y] = 0$$
$$E[xy] = E[x]E[y] = 0$$

$$\text{var}(x+y) = E[(x+y)^2] = E[x^2 + 2xy + y^2]$$

$$= E[x^2] + 2E[xy] + E[y^2] = \text{var}(x) + \text{var}(y)$$

- Examples:

  - If $X = Y$: $\text{var}(X + Y) = \text{var}(2x) = 4\text{var}(x)$

  - If $X = -Y$: $\text{var}(X + Y) = \text{var}(0) = 0$

  - If X, Y independent: $\text{var}(X - 3Y) = \text{var}(x) + \text{var}(-3y) = \frac{\text{var}(x)}{+9\text{var}(y)}$

Variance of the binomial

- X: binomial with parameters n, p

  − number of successes in n independent trials

$X_i = 1$ if ith trial is a success;   (indicator variable)   *independent*
$X_i = 0$ otherwise

$X = X_1 + \cdots + X_n$

$$\boxed{\text{var}(x)} = var(X_1) + \cdots + Var(X_n)$$

$$= n \cdot Var(X_1) = \boxed{np(1-p)}$$

The hat problem

- n people throw their hats in a box and then pick one at random
  - All permutations equally likely $1/n!$
  - Equivalent to picking one hat at a time

- X: number of people who get their own hat
  - Find $E[X]$ $= E[X_1] + \cdots + E[X_n] = n \cdot \frac{1}{n} = \boxed{1}$

$$X_i = \begin{cases} 1, & \text{if } i \text{ selects own hat} \\ 0, & \text{otherwise.} \end{cases}$$

$$X = X_1 + X_2 + \cdots + X_n$$

- $E[X_i] = E[X_1] = P(X_1 = 1) = \frac{1}{n}$

people   hats

2   [1] ⊠

1   [2] ⊠ $X = 1$

3   [3] ———— [3]

$\frac{1}{3} \cdot \frac{1}{2} \cdot 1 = \frac{1}{3!}$

$P_x(k):$ hard!

$\sum_k k \, P_x(k)$

The variance in the hat problem

$n = 2$

$X_1 = 1 \implies X_2 = 1$

$X_1 = 0 \implies X_2 = 0$

- X: number of people who get their own hat

  – Find var(X)

$$X_i = \begin{cases} 1, & \text{if } i \text{ selects own hat} \\ 0, & \text{otherwise.} \end{cases}$$

$X = X_1 + X_2 + \cdots + X_n$

$\frac{n(n-1)}{n^2 - n}$

- $\boxed{\text{var}(X)} = \mathbf{E}[X^2] - (\mathbf{E}[X])^2 \approx 2 - 1 = \boxed{1}$

$$X^2 = \overbrace{\sum_i X_i^2}^{n} + \overbrace{\sum_{i,j:i\neq j} X_i X_j}^{n^2 - n}$$

- $\mathbf{E}[X_i^2] = E[X_1^2] = E[X_1] = 1/n$

  $E[X^2] = n \cdot \frac{1}{n} + n(n-1) \cdot \frac{1}{n} \cdot \frac{1}{n-1}$

- For $i \neq j$: $\mathbf{E}[X_i X_j] = E[X_1 X_2] = P(X_1 X_2 = 1) = P(X_1 = 1, X_2 = 1)$

  $= P(X_1 = 1) P(X_2 = 1 | X_1 = 1) = \frac{1}{n} \cdot \frac{1}{n-1}$