CMPSCI 240: Reasoning about Uncertainty

Lecture 12: Multiple Random Variables

Andrew McGregor

University of Massachusetts

Last Compiled: March 22, 2017

Outline

Expectation and Variance Review

- The expected value $E[X]$ of a random variable X is a probability-weighted average of the possible values of X:

$$E[X] = \sum_k k\, P(X = k)$$

- If X is a random variable and $f : \mathbb{R} \to \mathbb{R}$ then $Y = f(X)$ is also a random variable with expectation

$$E(Y) = \sum_k f(k) P(X = k)$$

- The variance is quantifies how close to $\mu = E[X]$ we expect X to be:

$$\text{var}(X) = E[(X - \mu)^2] = E[X^2] - \mu^2.$$

- The standard deviation of X is $\sigma_X = \sqrt{\text{var}(X)}$

Markov and Chebyshev Bounds

- Markov Bound: For an non-negative random variable X,

$$P(X \geq c) \leq \frac{E(X)}{c}$$

- Chebyshev Bound: For a random variable X,

$$P(|X - E(X)| \geq c) \leq \frac{Var(X)}{c^2}$$

Multiple Random Variables

- Consider two random variables, X and Y mapping from Ω to $\mathbb{R}$.
- For $i, j \in \mathbb{R}$, we can define the event

$$\{X = i, Y = j\} = \{X{=}i\}\cap\{Y{=}j\} = \{o \in \Omega \mid X(o){=}i \text{ and } Y(o){=}j\}$$

- The probabilities of these events give the joint PMF of X and Y:

$$P(X = i, Y = j) = P(\{X = i, Y = j\})$$

Tabular Representation of Joint PMFs

P(X,Y)				
X\Y	$Y=1$	$Y=2$	$Y=3$	$Y=4$
$X=1$	0.1	0.1	0	0.2
$X=2$	0.05	0.05	0.1	0
$X=3$	0	0.1	0.2	0.1

- e.g., $P(X{=}2, Y{=}3) = 0.1$, $P(X{=}3, Y{=}1) = 0$, ...
- Given the joint PMF, how can we work out $P(X = i)$ and $P(Y = j)$?

$$P(X = i) = \sum_j P(X = i, Y = j)$$

$$P(Y = j) = \sum_i P(X = i, Y = j)$$

- If we start with the joint PMF of X and Y, we refer to $P(X)$ as the marginal PMF of X and $P(Y)$ as the marginal PMF of Y.

Outline

Conditioning

- Conditional PMF of X given Y:

$$P(X=i|Y=j) = P(\{X=i\}\,|\,\{Y=j\})\ .$$

- Compute $P(X|Y)$ using the definition of conditional probability:

$$P(X=i|Y=j) = \frac{P(X=i, Y=j)}{P(Y=j)}$$

  since for any two events A, B we have $P(A\,|\,B) = \frac{P(A\cap B)}{P(B)}$.

- The conditional probability $P(X=i|Y=j)$ is the joint probability $P(X=i, Y=j)$ normalized by the marginal $P(Y=j)$.
- An equivalent definition of independence is X and Y are independent if

$$\text{for all } i, j, \quad P(X=i|Y=j) = P(X=i)$$

Conditional PMFs

P(X,Y)				
X\Y	1	2	3	4
1	0.1	0.1	0	0.2
2	0.05	0.05	0.1	0
3	0	0.1	0.2	0.1

Y	1	2	3	4
P(Y)	0.15	0.25	0.3	0.3

$P(X\|Y)$				
X \Y	1	2	3	4
1	0.66	0.4	0	0.66
2	0.33	0.2	0.33	0
3	0	0.4	0.66	0.33

Outline

Functions of Two Random Variables

- Given two random variables X and Y and a function $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$,

$$Z = f(X, Y)$$

is a new random variable where

$$E(Z) = \sum_{x,y} f(x, y) P(X = x, Y = y) \text{ and } var(Z) = E(Z^2) - (E(Z))^2 .$$

- Linearity of Expectation: If $Z = X + Y$,

$$E(Z) = E(X + Y) = E(X) + E(Y)$$

- Linearity of Variance: If $Z = X + Y$,

$$var(Z) = var(X + Y) = var(X) + var(Y)$$

if X and Y are independent, i.e., for all i, j

$$P(X = i, Y = j) = P(X = i)P(Y = j) .$$

Linearity of Expectation

- **Lemma:** Given two random variables X, Y, and $Z = X + Y$ then

$$E[Z] = E[X + Y] = E[X] + E[Y]$$

- **Proof:** Generalized expected value rule.

$$\begin{aligned}
E[Z] &= \sum_a \sum_b (a+b) \cdot P(X=a, Y=b) \\
&= \sum_a \sum_b a \cdot P(X=a, Y=b) + \sum_a \sum_b b \cdot P(X=a, Y=b) \\
&= \sum_a a \sum_b P(X=a, Y=b) + \sum_b b \sum_a P(X=a, Y=b) \\
&= \sum_a aP(X=a) + \sum_b bP(Y=b) = E(X) + E(Y)
\end{aligned}$$

Expectation of Products of Independent Variables

- **Lemma:** If X and Y are independent then $E[XY] = E[X]E[Y]$:
- **Proof:**

$$
\begin{aligned}
E[XY] &= \sum_a \sum_b ab \cdot P(X = a, Y = b) \\
&= \sum_a \sum_b ab \cdot P(X = a)P(Y = b) \\
&= \sum_a a \cdot P(X = a) \cdot \sum_b b \cdot P(Y = b) \\
&= E[X]E[Y]
\end{aligned}
$$

Variance of Sums of Random Variables

- **Lemma:** If X and Y are independent then

$$var(X+Y) = var(X) + var(Y)$$

- **Proof:**

$$\begin{aligned}
var(X+Y) &= E[(X+Y)^2] - E[X+Y]^2 \\
&= E[X^2 + 2XY + Y^2] - (E[X] + E[Y])^2 \\
&= E[X^2] + 2E[X]E[Y] + E[Y^2] \\
&\quad - (E[X]^2 + 2E[X]E[Y] + E[Y]^2) \\
&= E[X^2] - E[X]^2 + E[Y^2] - E[Y]^2 \\
&= var(X) + var(Y)
\end{aligned}$$

Functions of Multiple Random Variables

- Given random variables $X_1, X_2, \dots, X_N$ and a function $f : \mathbb{R} \times \mathbb{R} \times \ldots \times \mathbb{R} \to \mathbb{R}$,

$$Z = f(X_1, X_2, \dots, X_N)$$

  is a new random variable.
- Linearity of Expectation: If $Z = \sum_{i=1}^{N} X_i$,

$$E(Z) = E(\sum_{i=1}^{N} X_i) = \sum_{i=1}^{N} E(X_i)$$

- Linearity of Variance: If $Z = \sum_{i=1}^{N} X_i$,

$$var(Z) = var(\sum_{i=1}^{N} X_i) = \sum_{i=1}^{N} var(X_i)$$

  if $X_1, \dots, X_N$ are pairwise independent, i.e., for all a, b

$$P(X_i = a, X_j = b) = P(X_i = a)P(X_j = b) \ .$$

Example 1

Toss 12 fair six-sided dice. Let X be the number of "1"s and let Y be the number of "6"s. Are X and Y independent:

A: Yes

B: No

C: Can't tell from the information given.

Answer is B since $P(X = 12, Y = 12) = 0 \neq P(X = 12)P(Y = 12)$.

Example 2

Toss 12 fair six-sided dice. Let X be the number of "1"s and let Y be the number of "6"s. What is the expected value of X?

A: 0

B: 1

C: 2

D: 3

E: 6

Answer is C since each of the 12 throws has a $1/6$ of being a "1".

Example 3

Toss 12 fair six-sided dice. Let X be the number of “1”s and let Y be the number of “6”s. What is the expected value of $X + Y$?

A: 0

B: 2

C: 4

D: 6

E: 12

Answer is C because $E(X + Y) = E(X) + E(Y) = 2 + 2 = 4$.

Outline

Secrets of the Markov Bound

- Markov Bound: For any non-negative random variable,

$$P(X \geq c) \leq E(X)/c$$

- For example, if $E(X) = 10$,

$$P(X \geq 15) \leq 2/3$$

- Can we infer an interesting upper bound about a lower tail, e.g.,

$$P(X \leq 5) \leq ???$$

- No! Consider a random variable of the form $X = 10t$ with probability $1/t$ and $X = 0$ with probability $1 - 1/t$. Then, $E(X) = 10$ but

$$P(X \leq 5) = 1 - 1/t$$

  can be arbitrarily close to 1.

Secrets of the Markov Bound 2

- We can do a bit better if we also know the maximum value of X.
- For example, suppose we know $0 \leq X \leq 15$ and $E(X) = 10$.
- Let $Y = 15 - X$ and note that

$$E(Y) = 15 - E(X) = 5$$

and $Y \geq 0$. Then,

$$P(X \leq 5) = P(Y \geq 10) \leq E(Y)/10 = 1/2$$

Secrets of the Chebyshev Bound

- Chebyshev Bound: $P(|X - E(X)| \geq c) \leq Var(X)/c^2$
- The bound is useful when we are trying to bound the probability that X is much smaller or larger than it's expectation.
- However, it also implies bounds on just one tail.
- For example, if $E(X) = 10$ and $var(X) = 2$ then

$$P(X \geq 15) = P(X \geq E(X) + 5) \leq P(|X - E(X)| \geq 5) \leq 2/25$$