Lecture 11
Random variables

CHOONG-WAN WOO | COCOAN lab | http://cocoanlab.github.io

Review: Key Points

Chapter 14: Randomness and Probability

- Terms: Trial, outcome/event, sample space ($\mathbf{S}$)
- Law of large numbers (LLN)
- Five basic rules of probability: $0 \leq P(\mathbf{A}) \leq 1$, $P(\boldsymbol{S}) = 1$, $P(\mathbf{A}^{C}) = 1 - P(\mathbf{A})$,

  $P(\mathbf{A}\ or\ \mathbf{B}) = P(\mathbf{A}) + P(\mathbf{B})$, when A and B are disjoint (or mutually exclusive),

  $P(\mathbf{A}\ and\ \mathbf{B}) = P(\mathbf{A}) \times P(\mathbf{B})$, when A and B are independent.

Chapter 15: Probability rules

- General addition rule: $P(\mathbf{A}\ or\ \mathbf{B}) = P(\mathbf{A}) + P(\mathbf{B}) - P(\mathbf{A}\ and\ \mathbf{B})$
- General multiplication rule: $P(\mathbf{A}\ and\ \mathbf{B}) = P(\mathbf{A}) \times P(\mathbf{B}|\mathbf{A}) = P(\mathbf{B}) \times P(\mathbf{A}|\mathbf{B})$
- Independence: $P(\mathbf{B}|\mathbf{A}) = P(\mathbf{B})$
- Bayes' Rule: $P(\mathbf{B}|\mathbf{A}) = \dfrac{P(\mathbf{A}|\mathbf{B})P(\mathbf{B})}{P(\mathbf{A}|\mathbf{B})P(\mathbf{B}) + P(\mathbf{A}|\mathbf{B}^{\mathbf{C}})P(\mathbf{B}^{\mathbf{C}})}$

Random Variables

- When its values are based on the outcome of a random event
- We denote random variables using a capital letter, like *X*
- If we can list all the outcomes, it's **discrete** random variable.
- Otherwise, it's a **continuous** random variable.

Policyholder Outcome	Payout x	Probability $P(X = x)$
Death	10,000	$\frac{1}{1000}$
Disability	5000	$\frac{2}{1000}$
Neither	0	$\frac{997}{1000}$

- Example of an insurance company
- Each year, the probability of death (death rate) is 1 out of every 1000 people, etc.
- We can't predict what will happen during any given year,
-     but we can say what we can *expect* to happen.
- What's the expected value of a policy payout?
- E(X) for expected value,
-     and we can use the mean (μ) to estimate it.

$$\begin{aligned}\mu &= E(X)\\ &= \$10{,}000\left(\frac{1}{1000}\right) + \$5000\left(\frac{2}{1000}\right) + \$0\left(\frac{997}{1000}\right)\\ &= \$20.\end{aligned}$$

$$\mu = E(X) = \sum xP(x)$$

(for discrete random variables)

Is this based on data?

- Yes, and no.
- Mean for data: $\bar{y} = \frac{Total}{n} = \frac{\sum y}{n}$
- Mean for random variables:

$$\mu = E(X) = \sum xP(x)$$

- What's differences?

- Probability conveys the information about **population**.
- Remember the law of large numbers.. The probability assumes a large number of repeats.

Law of Large Numbers (LLN)

- When we repeat a random process over and over, the proportion of times that an event occurs settle down to one number, which is the **probability** of the event.

Spread: Standard deviation

- Similar to the data case, we first calculate deviation from the mean and square it.
- Example of the insurance company again:

Policyholder Outcome	Payout x	Probability $P(X = x)$	Deviation $(x - \mu)$
Death	10,000	$\frac{1}{1000}$	$(10{,}000 - 20) = 9980$
Disability	5000	$\frac{2}{1000}$	$(5000 - 20) = 4980$
Neither	0	$\frac{997}{1000}$	$(0 - 20) = -20$

- The variance is the expected value of those squared deviations:

$$Var(X) = 9980^2\left(\frac{1}{1000}\right) + 4980^2\left(\frac{2}{1000}\right) + (-20)^2\left(\frac{997}{1000}\right) = 149{,}600.$$

- Its square root is standard deviation:

$$SD(X) = \sqrt{149{,}600} \approx \$386.78.$$

$$\sigma^2 = Var(X) = \sum (x - \mu)^2 P(x)$$
$$\sigma = SD(X) = \sqrt{Var(X)}$$

Shifting and combining random variables

- $E(X \pm c) = E(X) \pm c,\ Var(X \pm c) = Var(X)$
- $E(aX) = aE(X),\ Var(aX) = a^2Var(X)$
- $E(X \pm Y) = E(X) \pm E(Y)$
- If two random variables are independent, $Var(X \pm Y) = Var(X) + Var(Y)$
  - If they are not independent, $Var(X \pm Y) = Var(X) + Var(Y) \pm 2Cov(X,Y)$

Reminder

$$\mu = E(X) = \sum xP(x)$$

- $E(X \pm c) = E(X) \pm c,\ Var(X \pm c) = Var(X)$
- $E(aX) = aE(X),\ Var(aX) = a^2Var(X)$
- $E(X \pm Y) = E(X) \pm E(Y)$
- If two random variables are independent, $Var(X \pm Y) = Var(X) + Var(Y)$

Proof

$$\begin{aligned} E(X+Y) &= \sum_x \sum_y (x+y)P_{XY}(x,y) \\ &= \sum_x \sum_y xP_{XY}(x,y) + \sum_y \sum_x yP_{XY}(x,y) \\ &= \sum_x xP_X(x) + \sum_y yP_Y(y) \\ &= E(X) + E(Y) \end{aligned}$$

Covariance and correlation

- Remember the correlation between two variables from a previous chapter

$$r = \frac{\sum z_x z_y}{n-1} \qquad r = \frac{\sum (x-\bar{x})(y-\bar{y})}{\sqrt{\sum (x-\bar{x})^2 \sum (y-\bar{y})^2}} = \frac{\sum (x-\bar{x})(y-\bar{y})}{(n-1)s_x s_y}$$

- Now, correlation between two random variables.

Covariance and correlation

- Covariance between X and Y, where $E(X) = \mu, E(Y) = \nu,$
  - $Cov(X,Y) = E\big((X-\mu)(Y-\nu)\big)$
- Some properties of covariance
  1. $Cov(X,Y) = Cov(Y,X)$
  2. $Cov(X,X) = Var(X)$
  3. $Cov(cX, dY) = cdCov(X,Y)$, for any constants *c* and *d*
  4. $Cov(X,Y) = E(XY) - \mu\nu$
  5. If X and Y are independent, $Cov(X,Y) = 0$
     - but the converse is not always true
  6. $Var(X \pm Y) = Var(X) + Var(Y) \pm 2Cov(X,Y)$

Proofs

2. $\mathrm{Var}(X) = E((X-\mu_X)^2) = \mathrm{Cov}(X,X)$

4.
$$\begin{aligned} &\mathrm{Cov}(X,Y) \\ =\ & E((X-\mu_X)(Y-\mu_Y)) \\ =\ & E(XY - \mu_X Y - X\mu_Y + \mu_X\mu_Y) \\ =\ & E(XY) - \mu_X E(Y) - E(X)\mu_Y + \mu_X\mu_Y \\ =\ & E(XY) - \mu_X\mu_Y \end{aligned}$$

5.
$$\begin{aligned} \mathrm{Cov}(X,Y) &= E(XY) - \mu_X\mu_Y \\ &= E(X)E(Y) - \mu_X\mu_Y = 0 \end{aligned}$$

6.
$$\begin{aligned} & Var(X+Y) = E[(X+Y-\mu_x-\mu_y)^2] \\ & = E[(X-\mu_x)^2 + (Y-\mu_y)^2 + 2(X-\mu_x)(Y-\mu_y)] \\ & = Var(X) + Var(Y) + 2Cov(X,Y) \end{aligned}$$

Covariance and correlation

- Correlation: $Corr(X, Y) = \dfrac{Cov(X, Y)}{\sigma_X \sigma_Y}$

- The properties of the correlation:

Properties of Correlation

- **Sign of correlation:** the direction of the association (e.g., positive, negative)
- **Range:** *r* is always between -1 and 1.
  - When $r = 1$ all of the points lie on a straight line with a positive slope.
  - $r < 0$ indicates a negative association.
  - When $r = -1$ all points lie on a straight line with negative slope.
  - If *r* is close to 0, this indicates a very weak linear relationship.
- **Symmetry:** The correlation of x with y is the same as the correlation of y with x.
- **No units**
  - The value of *r* does not change even if units of measure are changed.
  - The correlation has no unit of measurement.
- **Only linear:** Correlation measures only the strength of a *linear* relationship.
- **Sensitive to outliers:** The correlation is sensitive to outliers.

Key Points

Chapter 16: Random variables

- Discrete vs. continuous random variables
- Expected values (mean): $\mu = E(X) = \sum xP(x)$
- Here, *probability* conveys the information about population assuming a large number of repeats
- Spread: $\sigma^2 = Var(X) = \sum (x - \mu)^2 P(x)$
  $\sigma = SD(X) = \sqrt{Var(X)}$
- $E(X \pm c) = E(X) \pm c,\ Var(X \pm c) = Var(X)$
- $E(aX) = aE(X),\ Var(aX) = a^2 Var(X)$
- $E(X \pm Y) = E(X) \pm E(Y)$
- $Cov(X, Y) = E\big((X - \mu)(Y - \nu)\big)$
- $Var(X \pm Y) = Var(X) + Var(Y) \pm 2Cov(X, Y)$
- $Corr(X, Y) = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$