Lecture 10 Notes

• Midterm
  – Good job overall! \( \mu = 81; \sigma = 18; \) Range: 30 – 110
  – Scores are up on TED
    • Also, updated Quiz scores including Week 4
    • If you are somehow not able to see your grades on TED, please email ABK + David + Pete + Natalie
  – Hand-back process
    • (1) Look at solutions + grading rubric on the class website
    • (2) Go to TA OH’s to see your MT
    • (3) If you’re satisfied, take your MT
    • (4) Else, leave your MT with a specific regrade request before 5/13
      • If you take away your MT, you cannot request a regrade later
      • All remaining MTs will be handed back at end of 5/13 lecture

• Goals for this week
  – Linearity of expectation; Joint distributions
  – Decision trees, hashing example
  – For quiz: MT solutions, mean/variance, today’s material
  – (Next week: conditional probability, Bayes’ rule; induction and recursion)
  • (Have caught up again to poker chip summary)
Expectation, Variance

- Random variable $X = \#\text{heads seen in four tosses of a fair coin}$

- **Expectation** $E(X) = \sum_{k=0,\ldots,4} p_k x_k =$

  $p_0 = P(X = 0) = 1/16 \quad * \quad 0 = 0$
  $p_1 = P(X = 1) = 4/16 \quad * \quad 1 = 4/16$
  $p_2 = P(X = 2) = 6/16 \quad * \quad 2 = 12/16$
  $p_3 = P(X = 3) = 4/16 \quad * \quad 3 = 12/16$
  $p_4 = P(X = 4) = 1/16 \quad * \quad 4 = 4/16$

$\sum = 32/16 = 2$
Expectation, Variance  Def. 8, FN-24

• Random variable $X = \#\text{heads seen in four tosses of a fair coin}$

- Variance $\sigma_X^2 = E(X^2) - (E(X))^2 = 5 - 2^2 = 1$

\[
\begin{align*}
p_0 &= P(X = 0) = 1/16 & \quad \ast \quad 0^2 &= 0 \\
p_1 &= P(X = 1) = 4/16 & \quad \ast \quad 1^2 &= 4/16 \\
p_2 &= P(X = 2) = 6/16 & \quad \ast \quad 2^2 &= 24/16 \\
p_3 &= P(X = 3) = 4/16 & \quad \ast \quad 3^2 &= 36/16 \\
p_4 &= P(X = 4) = 1/16 & \quad \ast \quad 4^2 &= 16/16
\end{align*}
\]
\[
\frac{80}{16} = 5
\]
FN Theorem 3: Linearity of Expectation

• If \( X \) and \( Y \) are random variables on the same sample space and if \( a \) and \( b \) are real numbers, then:

\[
E(aX) = aE(X)
\]

\[
E(aX + bY) = aE(X) + bE(Y)
\]

Holds regardless of whether \( X \) and \( Y \) are independent

• Example: Roll two dice. What is the expected value of their sum?
Expected Sum of Two Rolled Dice

• Define r.v.’s:
  – $X =$ value of first die’s roll
  – $Y =$ value of second die’s roll
  – $Z =$ sum of two values: $Z = X + Y$

• Linearity of expectation $\Rightarrow E(Z) = E(X) + E(Y)$
Thank-You Notes

- A child has just had a birthday party. He has written 20 thank-you notes and addressed 20 corresponding envelopes. But then he gets distracted, and puts the letters randomly into the envelopes (one letter per envelope). What is the expected number of letters that are in their correct envelopes?

Define r.v. \( X = \# \text{letters in correct envelopes} \)?

\[
E(X) = \sum_{i=0}^{20} i \cdot P(X = i)
\]

- How about: r.v.’s \( X_j = 1 \) if letter \( j \) is in envelope \( j \), and \( X_j = 0 \) otherwise?

\[
E(\sum_{j=1}^{20} X_j)
\]
Joint Distribution

- When two random variables are defined on the same sample space, can consider their **joint distribution function**, $h$:

  - $h_{X,Y} : \text{Im}(X) \times \text{Im}(Y) \rightarrow [0,1]$
  
  - $h_{X,Y}(i,j) = P(X^{-1}(i) \cap Y^{-1}(j))$

    probability of ((X = i) AND (Y = j))
Joint Distribution

• Recall: Flip a fair coin four times.
  \[ U = \{(f_1,f_2,f_3,f_4) \mid f_i \in \{H,T\}\} \]

  – r.v. \( X_1 \) = number of heads that appear
  – r.v. \( X_2 \) = 1 if the second flip is heads, 0 otherwise
  – r.v. \( X_3 \) = 1 if the third flip is heads, 0 otherwise

• What is the joint distribution \( h_{X_1,X_2} \) of \( X_1 \) and \( X_2 \)?

• Domain of \( h_{X_1,X_2} \) = \{\( (i,j) \mid i \in \text{Im}(X_1), j \in \text{Im}(X_2) \}\)
  = \{0,1,2,3,4\} \times \{0,1\}
Joint Distribution

• **Recall:** Flip a fair coin four times.
  - r.v. $X_1 =$ number of heads that appear
  - r.v. $X_2 = 1$ if the second flip is heads, 0 otherwise

• What is $h_{X_1,X_2}(2,1)$?

• $h_{X_1,X_2}(2,1)$ is a probability …. of what event?

• $h_{X_1,X_2}(2,1) = P(\{(f_1,f_2,f_3,f_4) \mid \text{two H’s, second toss is H}\})$
  $= P(\{(f_1,H,f_3,f_4) \mid \text{one of } f_1,f_3,f_4 \text{ is H}\})$ $C(3,1)2^{-4}$
  $= P(\{(H,H,H,T), (T,H,H,T), (T,H,T,H)\})$
  $= 3/16$
Joint Distribution

• Recall: Flip a fair coin four times.
  – r.v. $X_1 =$ number of heads that appear
  – r.v. $X_2 = 1$ if the second flip is heads, 0 otherwise

• What is $h_{X_1,X_2} (0,1)$?
  – $h_{X_1,X_2} (0,1) = P\{((f_1,f_2,f_3,f_4) | \text{zero H's, second toss is H})
  \quad = P(\emptyset) = 0$

• What is $h_{X_1,X_2} (4,0)$?
  – $h_{X_1,X_2} (4,0) = P\{((f_1,f_2,f_3,f_4) | \text{four H's, second toss is T})

• $h_{X_1,X_2} (3,1)$?

• $h_{X_1,X_2} (3,0)$?

• The joint distribution $h_{X_2,X_3}$? (exercise)
Independence  Definition 10, FN-29/30

• X and Y are independent random variables if and only if \( h_{X,Y}(i,j) = f_X(i) \cdot f_Y(j) \) for all \((i,j) \in \text{Im}(X) \times \text{Im}(Y)\)

• Intuitively: information about one of the r.v.’s doesn’t give us any information about the other

• Yet again: Flip a fair coin four times.
  – r.v. \( X_1 = \) number of heads that appear
  – r.v. \( X_2 = 1 \) if the second flip is heads, 0 otherwise
  – r.v. \( X_3 = 1 \) if the third flip is heads, 0 otherwise

• Do you think that \( X_1 \) and \( X_2 \) are independent?
• Do you think that \( X_2 \) and \( X_3 \) are independent?
Independence  Definition 10, FN-29/30

• X and Y are independent random variables if and only if \( h_{X,Y}(i,j) = f_X(i) \cdot f_Y(j) \) for all \((i,j) \in \text{Im}(X) \times \text{Im}(Y)\)

  – r.v. \( X_1 \) = number of heads that appear
  – r.v. \( X_2 \) = 1 if the second flip is heads, 0 otherwise
  – r.v. \( X_3 \) = 1 if the third flip is heads, 0 otherwise

• **How can we show \( X_1 \) and \( X_2 \) are not independent?**

  \[
  h_{X_1,X_2}(2,1) = \qquad f_{X_1}(2) = \qquad f_{X_2}(1) = \\
  h_{X_1,X_2}(1,1) = \qquad f_{X_1}(1) = \qquad f_{X_2}(1) = 
  \]
Independence  Definition 10, FN-29/30

• X and Y are independent random variables if and only if \( h_{X,Y}(i,j) = f_X(i) \cdot f_Y(j) \) for all \((i,j) \in \text{Im}(X) \times \text{Im}(Y)\)

  – r.v. \( X_1 \) = number of heads that appear
  – r.v. \( X_2 \) = 1 if the second flip is heads, 0 otherwise
  – r.v. \( X_3 \) = 1 if the third flip is heads, 0 otherwise

• How can we show \( X_2 \) and \( X_3 \) are independent?

  \[
  h_{X_2,X_3}(0,0) = \quad f_{X_2}(0) = \quad f_{X_3}(0) = \\
  h_{X_2,X_3}(1,1) = \quad f_{X_1}(1) = \quad f_{X_2}(1) = \\
  \]

  \[... |\text{Im}(X_2) \times \text{Im}(X_3)| = 4 \Rightarrow \text{this can be tedious!} \]
Independence  Definition 10, FN-29/30

• If X and Y are independent random variables and f, g : R → R are functions, then f(X), g(Y) are independent random variables.

• \( E(XY) = E(X) \cdot E(Y) \).

  not true if the r.v.’s are not independent!
Independence  Definition 10, FN-29/30

- $E(X \cdot Y) = E(X) \cdot E(Y)$.
  
  not true if the r.v.'s are not independent!

- Coin flipping:
  
  - $E(X_1) = 2, \quad E(X_2) = \frac{1}{2}$
  
  - $Y = X_1 \cdot X_2 \in \{i \cdot j : i \in \text{Im}(X_1), j \in \text{Im}(X_2)\} = \{0,1,2,3,4\}$
  
  - $E(X_1 \cdot X_2) = \sum_{k=0,\ldots,4} P(Y = k) \cdot k$

  $$
  = \frac{1}{16} \cdot 1 \\
  + \frac{3}{16} \cdot 2 \\
  + \frac{3}{16} \cdot 3 \\
  + \frac{1}{16} \cdot 4
  $$
Independence  Definition 10, FN-29/30

- $E(\cdot \cdot) = E(\cdot) \cdot E(\cdot)$.
  not true if the r.v.’s are not independent!

- Coin flipping:
  - $E(X_2) = \frac{1}{2}$,  $E(X_3) = \frac{1}{2}$
  - $Y = X_2 \cdot X_3 \in \{i \cdot j : i \in \text{Im}(X_2), j \in \text{Im}(X_3)\} = \{0, 1\}$
  - $E(X_2 \cdot X_3) = \sum_{k=0,1} P(Y = k) \cdot k$

  $= P(X_2 \cdot X_3 = 0) \cdot 0$
  $+ P(X_2 \cdot X_3 = 1) \cdot 1$
Decision Trees

- Decision trees are a systematic way of listing templates / possibilities
  - Decision trees are a type of graph
  - Can be used for enumeration
  - Can be used to capture recursive solution approaches (algorithms)
  - Vehicle for correspondences between induction and recursion
Decision Trees

• Example: How many functions from $A = \{1,2,3,4\}$ to itself are non-increasing?
• Each function can be defined by a sequence of decisions:
  
  What is $f(1)$
  What is $f(2)$
  What is $f(3)$
  What is $f(4)$ …
  … but there are not independent decisions!

  – If $f(1) = 1$, then we must have $f(2) = f(3) = f(4) = 1$
  – If $f(1) = 2$, then we must have $f(2) \in \{1,2\}$
  – If $f(2) = 2$, then $f(3) \in \{1,2\}$
  – Etc.

  (There are 35 such functions)
Decision Tree: Non-Attacking Queens

- Put N queens on an N x N chessboard such that no queen attacks another queen or, “in how many ways can …”

A queen in chess can attack any square on the same row, column, or diagonal. Given an n x n chessboard, we seek to place n queens onto squares of the chessboard, such that no queen attacks another queen. The example shows a placement (red squares) of four mutually non-attacking queens.
Non-Attacking Queens: Leaf in Tree = Solution
Non-Attacking Queens Decision Tree

Comments

• Our tree corresponded to a solution space of size $4^4 = 256$

• What if we numbered the squares as 1, 2, …, 16 and tried to enumerate all $\binom{16}{4}$ choices?
  – Our solution space would have size 1820

• Our decision tree was “smarter”
  – We realized that there must be exactly one queen per row, so the decision for the $k^{th}$ row (i.e., $k^{th}$ queen) boils down to picking a square in the row

• Our illustration of the tree did not show any infeasible configurations
A graph is legally colored with k colors if each vertex has a color (label) between 1 and k (inclusive), and no adjacent vertices have the same color.

Example application: scheduling classrooms
A graph is legally colored with k colors if each vertex has a color (label) between 1 and k (inclusive), and no adjacent vertices have the same color.

Does the ordering of the vertices (1, 2, 3, 4, 5) matter?
Hashing

Example: Hashing application

- Definition: A hash function is any algorithm that maps data of variable length to data of a fixed length. The values returned by a hash function are called hash values, hash codes, hash sums, checksums or simply hashes.

- Open Hashing (closed addressing): Each item hashes to a particular location in an array, and locations can hold more than one item. However, multiple input data values can have the same hash value. These are called collisions!
Hashing

• What is the expected number of items per location?

• Start with $n$ items, hash to table of size $k$

• Can model as $n$ independent trials with outcome being the location in table (i.e., hash value)

• Let r.v. $X =$ number of inputs that hash to location 1
  
  - Can write this as the sum of:
    
    • $X_i = 1$ if the $i^{th}$ input hashes to location 1
    • $X_i = 0$ otherwise
  
  $\Rightarrow X = X_1 + X_2 + \ldots + X_n$

  $\Rightarrow E(X) = E(X_1) + E(X_2) + \ldots + E(X_n)$

  $= n \cdot 1/k = n/k$
Hashing (cont.)

• What is the expected number of empty locations?

• If I hash three keys into a table with 10 slots, what is the probability that all the keys go to different slots?
Problems 10

• P10.1 To receive credit, you must solve both (a) and (b). (a) Consider a row of \( n \) seats. A child sits on each. Each child may move by at most one seat. Find the number of ways in which the children can rearrange. (b) Same as in (a), but the children sit in a circle of \( n \) seats.

• P10.2 Two fishing sinker manufacturers A and B together offer 2001 distinct sinkers (with all distinct weights) in their catalogs. Manufacturer A has sinkers labeled \( a_1, \ldots, a_{1000} \), with the corresponding weights being \( a_1 < a_2 < \ldots < a_{1000} \). Similarly, manufacturer B has sinkers labeled \( b_1, \ldots, b_{1001} \), with the corresponding weights being \( b_1 < b_2 < \ldots < b_{1001} \). You have one copy of each of these 2001 sinkers. Explain how to find the sinker whose weight is ranked 1001, using 11 weighings. (Each weighing on a balance can tell you whether the contents of the left pan are heavier or lighter than the contents of the right pan.)
Problems 10 (cont.)

• P10.3 Prove that you can choose at most \( \frac{2^n}{n+1} \) \( n \)-bit binary strings, such that any two of these strings differ in at least three bit positions. (A motivation for this would be error-tolerant codes.)

• 10.4 To receive credit, you must solve both (a) and (b). You have 128 objects with distinct weights. (a) Find, with explanation, the minimum number of pairwise comparisons needed to identify the heaviest and second heaviest of these objects. (b) Find, with explanation, the minimum number of pairwise comparisons needed to identify the heaviest and the lightest of these objects.