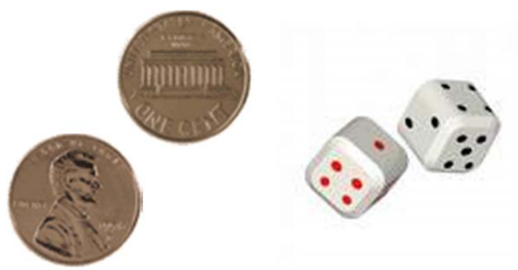
CMPS 2200 – Fall 2015



Probability and Expected Values

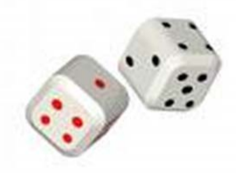
Carola Wenk

Probability

- Let S be a **sample space** of possible outcomes.
- $E \subseteq S$ is an **event**
- The (Laplacian) **probability of** E is defined as $P(E)=|E|/|S|$
  $\Rightarrow P(s)=1/|S|$ for all $s \in S$

Note: This is a special case of a probability distribution. In general $P(s)$ can be quite arbitrary. For a loaded die the probabilities could be for example $P(6)=1/2$ and $P(1)=P(2)=P(3)=P(4)=P(5)=1/10$.

Example: Rolling a (six-sided) die



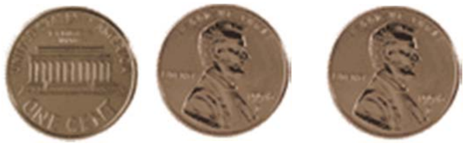
- $S = \{1,2,3,4,5,6\}$
- $P(2) = P(\{2\}) = 1/|S| = 1/6$
- Let $E = \{2,6\} \Rightarrow P(E) = 2/6 = 1/3 =$ P(rolling a 2 or a 6)

In general: For any $s \in S$ and any $E \subseteq S$

- $0 \leq P(s) \leq 1$
- $\sum_{s \in S} P(s) = 1$
- $P(E) = \sum_{s \in E} P(s)$

Random Variable

- A random variable X on S is a function from S to R
  $X: S \rightarrow \mathbf{R}$



Example 1: Flip coin three times.

- S = {HHH, HHT, HTH, HTT, THH, THT, TTH, TTT}
- Let $X(s)$ = # heads in s
  $\Rightarrow X(\text{HHH}) = 3$
  $X(\text{HHT})=X(\text{HTH})=X(\text{THH}) = 2$
  $X(\text{TTH})=X(\text{THT})=X(\text{HTH}) = 1$
  $X(\text{TTT}) = 0$

Example 2: Play game: Win \$5 when getting HHH, pay \$1 otherwise

- Let $Y(s)$ be the win/loss for the outcome s
  $\Rightarrow Y(\text{HHH}) = 5$
  $Y(\text{HHT}) = Y(\text{HTH}) = \ldots = -1$

What is the *average* win/loss?

Expected Value

- The **expected value** of a random variable $X\colon S \to \mathbf{R}$ is defined as
  $E(X) = \sum_{s \in S} P(s) \cdot X(s) = \sum_{x \in \mathbf{R}} P(\{X=x\}) \cdot x$

Notice the similarity to the **arithmetic mean (**or **average)**.

Example 2 (continued):

- $E(Y) = \sum_{s \in S} P(s) \cdot Y(s) = P(HHH) \cdot 5 + \sum_{s \in S \setminus \{HHH\}} P(s) \cdot (-1) = 1/2^3 \cdot 5 + 7 \cdot 1/2^3 \cdot (-1)$
  $= (5-7)/2^3 = -2/8 = -1/4$
  $= \sum_{x \in \mathbf{R}} P(\{Y=y\}) \cdot y = P(HHH) \cdot 5 + P(HHT) \cdot (-1) + P(HTH) \cdot (-1) + P(HTT) \cdot (-1) + P(THH) \cdot (-1) + P(THT) \cdot (-1) + P(TTH) \cdot (-1) + P(TTT) \cdot (-1)$

$\Rightarrow$ The average win/loss is $E(Y) = -1/4$

Theorem (Linearity of Expectation):

Let X, Y be two random variables on S. Then the following holds:

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

Proof: $E(X+Y) = \sum_{s \in S} P(s) \cdot (X(s)+Y(s)) = \sum_{s \in S} P(s) \cdot X(s) + \sum_{s \in S} P(s) \cdot Y(s) = E(X) + E(Y)$ □

Randomized algorithms

- Allow random choices during the algorithm
- Sample space S = {all sequences of random choices}
- The runtime $T: S \rightarrow \mathbf{R}$ is a random variable. The runtime $T(s)$ depends on the particular sequence s of random choices.

$\Rightarrow$ Consider the **expected runtime** $E(T)$