Math 2200-002/Discrete Mathematics

Combinatorics

Definition. (a) A *sample space* S is a set of equally likely outcomes.

(b) An *event* is a subset $E \subseteq S$ of the sample space.

(c) The *probability* of the event E is:

$$p(E) = \frac{|E|}{|S|}$$

Example. The probability of rolling a 7 with two die.

- S is the set of ordered pairs (i, j) with $1 \le i \le 6$ and $1 \le j \le 6$. (all the possible rolls of a pair of dice)
- E is the set $\{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$. Thus:

$$p(E) = \frac{|E|}{|S|} = \frac{6}{36} = \frac{1}{6}$$

Features of Probabilities.

- $0 \le p(E) \le 1$.
- $p(\overline{E}) = 1 - p(E)$ if $\overline{E}$ is the *complement* of E in S. (The probability of an event E **not** happening is $1 - p(E)$)
- $p(E \cup F) = p(E) + p(F) - p(E \cap F)$ if E and F are two events. (This follows from the inclusion-exclusion principle)

Example. The probability that a number chosen at random between 1 and 100 is divisible by 2 or 5 is:

$$p(\text{divisible by 2}) + p(\text{divisible by 5}) - p(\text{divisible by 2 and 5})$$

$$= \frac{1}{2} + \frac{1}{5} - \frac{1}{10} = \frac{3}{5}$$

Instead of each outcome being equally likely (which is not always realistic), we can **assign** probabilities to elements of the sample space. That is, we give probabilities:

$$0 \le p(s) \le 1 \text{ to each element } s \in S$$

such that:

$$\sum_{s \in S} p(s) = 1$$

This is called a **probability distribution**. In this setting, if $E \subseteq S$ is an event, then:

$$p(E) = \sum_{s \in E} p(s)$$

Example. "Equally likely outcomes" is the *uniform distribution*:

$$p(s) = \frac{1}{|S|} \text{ for each } s \in S$$

Example (Biased coin) Suppose a biased coin has probability p of landing heads and a probability q of landing tails with $p + q = 1$. If the coin is flipped n times, then:

S is the space of ordered n-tuples of binary choices between heads and tails, which we represent with digits 1 (for heads) and 0 (for tails). Thus, for example, two flips of the coin has a space S of four outcomes:

$$(0,0), (0,1), (1,0), (1,1)$$

Each outcome has its own probability. In this $n = 2$ example,

$$p(0,0) = q^2, p(0,1) = qp, p(1,0) = pq \text{ and } p(1,1) = p^2$$

(and note that these add to $p^2 + 2pq + q^2 = (p+q)^2 = 1$)

If E is the event that r heads and $n - r$ tails are flipped, then:

$|E| = C(n, r)$ is the number of positions for r heads among the n flips

$$\text{and } p(e) = p^r q^{n-r} \text{ for each element } e \in E$$

so putting these together:

$$p(E) = \sum_{e \in E} p^r q^{n-r} = C(n,r) p^r q^{n-r}$$

This Example is also known as a **Bernoulli trial**.

More Definitions. (a) A *random variable* is a function $X : S \to \mathbb{R}$.

(b) In the presence of a probability distribution, the *expected value* of a random variable X is:

$$E(X) = \sum_{s \in S} p(s) X(s) = \sum_{r \in \mathbb{R}} p(X = r) r$$

where $p(X = r) = \sum_{\{s | X(s) = r\}} p(s)$.

Example. Let S be the sample space of n flips of the biased coin and:

$$X(s) = \text{ the number of heads among the } n \text{ flips}$$

Then:

$$E(X) = \sum_{r=0}^{n} C(n,r) p^r q^{n-r} \cdot r = np$$

using the value $p(X = r) = C(n,r) p^r q^{n-r}$ that we found above.
(the surprising simplification of the sum is done in the book).

Definition. The *variance* of X with distribution $p(s)$ is:

$$V(X) = \sum_{s\in S}(X(s) - E(X))^2 p(s)$$

and the *standard deviation* is:

$$\sigma = \sqrt{V(X)}$$

This can be formidable to compute by hand when S is large.

Proposition. Let $X^2 : S \to \mathbb{R}$ be the function $X^2(s) = X(s)^2$. Then:

$$V(X) = E(X^2) - E(X)^2$$

i.e. the variance is the difference between the expected value of the square of the random variable and the square of the expected value.

Proof.

$$V(X) = \sum_{s\in S}(X(s) - E(X))^2 p(s) =$$

$$= \sum_{s\in S} X^2(s)p(s) - 2E(X)X(s)p(x) + E(X)^2 p(s)$$

$$= \sum_{s\in S} X^2(s)p(s) - 2E(X)\sum_{s\in S} X(s)p(s) + E(X)^2 \sum_{s\in S} p(s)$$

$$= E(X^2) - 2E(X)E(X) + E(X)^2 = E(X^2) - E(X)^2$$

Suppose now that X and Y are **two** random variables on S. Then:

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

Proof.

$$E(X+Y) = \sum_{s\in S}(X+Y)(s)p(s) = \sum_{s\in S}(X(s) + Y(s))p(s) =$$

$$= \sum_{s\in S} X(s)p(s) + \sum_{s\in S} Y(s)p(s) = E(X) + E(Y).$$

It is **not**, in general, true that $E(XY) = E(X)E(Y)$, but:

Definition. Random variables X and Y are **independent** if:

$$p(X = r_1 \text{ and } Y = r_2) = p(X = r_1) \cdot p(Y = r_2)$$

for all pairs of real numbers (r_1, r_2).

Examples. If a coin is tossed n times, then the random variables

$$X_i = \begin{cases} 1 \text{ if the } i\text{th flip is heads} \\ 0 \text{ if the } i\text{th flip is tails} \end{cases}$$

are independent. Similarly, if a pair of dice is rolled, then:

$$X_1(i,j) = i \text{ and } X_2(i,j) = j$$

are independent random variables.

On the other hand,

$$X(i,j) = i \text{ and } Y(i,j) = i + j$$

are not random variables, since, for example:

$$p(X = 1 \text{ and } Y = 8) = 0$$

but $p(X = 1) = 1/6$ and $p(Y = 8) = 5/36$.

Proposition. If X and Y are independent random variables, then:

$$E(XY) = E(X)E(Y)$$

Proof.

$$E(XY) = \sum_r p(XY = r)r = \sum_{r_1}\sum_{r_2} p(X = r_1 \text{ and } Y = r_2)r_1r_2 =$$

(summing over all the possible ways to factor $r = r_1r_2$)

$$= \sum_{r_1}\sum_{r_2} p(X = r_1)p(Y = r_2)r_1r_2 = \left(\sum_{r_1} p(X = r_1)r_1\right)\left(\sum_{r_2} p(X = r_2)r_2\right)$$

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

Notice. It was crucial to the proof that X and Y be independent!

Proposition. If X and Y are independent random variables, then:

$$V(X + Y) = V(X) + V(Y)$$

Proof.

$$V(X + Y) = E((X + Y)^2) - E(X + Y)^2 =$$

$$= E(X^2 + 2XY + Y^2) - \big(E(X)^2 + 2E(X)E(Y) + E(Y)^2\big)$$

$$= (E(X^2) - E(X)^2) + 2(E(XY) - E(X)E(Y)) + E(Y^2) - E(Y)^2$$

$$= V(X) + V(Y)$$

(the middle terms vanishes by the previous proposition!).

Examples. (a) Flip a fair coin n times. Then each of the random variables X_i defined above satisfies:

$$E(X_i) = \frac{1}{2} \text{ and } V(X_i) = E(X_i^2) - E(X_i)^2 = \frac{1}{2} - \frac{1}{4} = \frac{1}{4}$$

Thus, the expected value of:

$$X = X_1 + \cdots X_n = \text{total number of heads}$$

is $E(X) = n/2$ and the variance is $V(X) = n/4$.

(b) Roll a pair of dice and let $X_1(i,j) = i$ and $X_2(i,j) = j$. Then:

$$E(X_1) = E(X_2) = \frac{1}{6}(1+2+3+4+5+6+7) = \frac{7}{2}$$

and

$$E(X_1^2)^2 = E(X_2^2) = \frac{1}{6}(1+4+9+16+25+36) = \frac{91}{6}$$

so

$$V(X_1) = V(X_2) = \frac{91}{6} - \frac{49}{4} = \frac{35}{12}$$

from which it follows that:

$$E(X_1 + X_2) = 7 \text{ and } V(X_1 + X_2) = \frac{35}{6}$$

and $\sigma \sim 2.4$.

Informal Rule. There is an informal rule for random variables, which is:

$$p(|X - E(X)| < \sigma) \approx \frac{2}{3}$$

and

$$p(|X - E(X)| < 2\sigma) \approx 0.95$$

If we try this out on the pair of dice, we get:

$$p(|X - 7| < 2.4) = \frac{1}{36}(4+5+6+5+4) = \frac{24}{36} = \frac{2}{3}$$

(the probability of rolling a $5, 6, 7, 8$ or 9) and

$$p(|X - 7| < 4.8) = 1 - \left(\frac{1}{36}(1+1)\right) = \frac{34}{36} \approx 0.94$$

which is pretty much on the money for the informal rule!

Homework. (Due in class on Monday, December 2).

1. (2 points) What is the probability that a number chosen at random between 1 and 1000 is divisible by 7 or 11 or 13?

2. (4 points) Roll a pair of dodecahedral dice (12 sides). Then:

$$S = \{(i, j) \mid 1 \le i \le 12 \text{ and } 1 \le j \le 12\}$$

Imitate the exercises above with $X_1(i, j) = i$ and $X_2(i, j) = j$. That is:

(a) Find $E(X_i)$ and $V(X_i)$

(b) Find $E(X), V(X)$ and $\sigma = \sqrt{V(X)}$ for $X = X_1 + X_2$.

(c) Check the informal 2/3 and 95% rules.

3. (4 points) Take a biased coin with $p(\text{heads}) = 1/3$ and $p(\text{tails}) = 2/3$ and flip it 10 times.

(a) What is $|S|$?

Let X be the random variable $X(s) =$ number of heads.

(b) Calculate $E(X)$ and $V(X)$ and $\sigma = \sqrt{V(X)}$.

(c) Compute: $p(|X - E(X)|) < \sigma)$ and $p(|X - E(X)|) < 2\sigma$ and check that the informal rule sucks for biased coins. (The problem is with the probability distribution, which is skewed).

Hint: Use the fact that $X = X_1 + \cdots + X_{10}$ is the sum of 10 independent random variables!