STAT 270

Fall 2015

Richard Lockhart

Probability

Purposes of These Notes

- Jargon: experiment, sample space, outcome, event.

- Set theory ideas and notation: intersection, union, complement.

- Venn Diagrams

- Empty event or empty set.

- Disjoint, mutually exclusive.

- de Morgan's laws.

- Axioms of probability, probability measure, countable additivity.

- Symmetry, counting methods.

- Long-run relative frequency.

Sample Spaces: $\mathcal{S}$

- Informal idea: do "random" experiment.
- Outcome not certain beforehand.
- **Sample space** is list of all possible outcomes.
- Often idealized.
- **Example**: Toss coin 3 times, get H (heads) or T (tails) each time.
- Total of 8 possible outcomes in Sample Space.
- Richard lists $\mathcal{S}$ in class.

More sample spaces

- **Example**: Hang a 2 kg weight from a piece of piano wire.

- Measure how much wire stretches.

- Repeat with 4 kg, 6Kg, 8kg, 10kg.

- Sample space is

$$\mathcal{S} = \{(l_1, l_2, l_3, l_4, l_5)\}$$

  where the l_i are supposed to be any possible measured stretch.

- If using ruler marked to nearest millimetre then each l_i must be a multiple of 1 mm.

- Can the l_i be negative?

- Sample space can be chosen for convenience.

- Likely to choose $\mathcal{S} = \mathbb{R}^5$.

Events

- Technically: subsets of $\mathcal{S}$ for which we can calculate probabilities.

- **Example**: 3 coin tosses.

- First toss is Heads:

$$E = \{ \qquad\qquad\qquad\qquad\qquad\qquad \}$$

- Even number of heads:

$$E = \{ \qquad\qquad\qquad\qquad\qquad\qquad \}$$

- **Example**: piano wire (Hooke's law):

- Average is more than 10 mm:

$$E = \{ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \}$$

Set Notation

- Many probabilities will be calculated by breaking events up, putting events together, or working with opposites.

- *Complement* of event, A, is the event that A does not happen:

$$\bar{A} = \{s \in \mathcal{S} : s \notin A\}.$$

- *Intersection*: A, B events; event that both A and B happen:

$$AB = A \cap B = \{s \in \mathcal{S} : s \in A \text{ and } s \in B\}.$$

- *Union*: A, B events; event that at least one of A or B happens:

$$A \cup B = \{s \in \mathcal{S} : s \in A \text{ or } s \in B\}.$$

- Similarly for $A_1 \cup \cdots \cup A_n$ or $A_1 A_2 \cdots A_n$.

- In words: $A_1 \cup \cdots \cup A_n =$

  event that at least one of A_1 to A_p happens

  $A_1 A_2 \cdots A_n =$ all the events A_1 to A_p happen.

- Sometimes helpful to use Venn diagrams.

Venn Diagrams

Richard will draw some Venn diagrams.

Empty set, disjoint events, mutually exclusive

- The empty set is $\emptyset$. It has no elements.

- Always an event. Sometimes called *null event*.

- Two events, A and B are *disjoint* or *mutually exclusive* if

$$AB = \emptyset.$$

- A collection of events $A_1, A_2, \ldots$ is *mutually exclusive* or *pairwise disjoint* if, for each pair $i \neq j$, we have

$$A_i A_j = \emptyset.$$

- Get Richard to draw De Morgan's laws:

$$(A \cup B)' = A' \cap B'$$

and

$$(A \cap B)' = A' \cup B'$$

More Venn Diagrams

Richard will draw more Venn Diagrams.

Some English

- "No more than 1 head" same as "fewer than 2 heads" same as "less than or equal to 1 head" same as "not either 2 heads or 3 heads" same as "not at least two heads".

- All these are

$$\{HTT, THT, TTH, TTT\}.$$

- Question: Toss coin 4 times.

A is the event I get 3 heads.

How many *elements* in A? (What is the *cardinality* of A?)

Probability

- Each event A has a probability $P(A)$.
- Axioms of probability: Academician Kolmogorov, 1933.
- Probabilities: real numbers, can't be negative:

$$\forall A P(A) \geq 0.$$

- Events made by putting together mutually exclusive events have probabilities which add up:

If $\forall i \neq j$ we have $A_i \cap A_j = \emptyset$ then

$$P\left(\cup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

- Draw 3 mutually exclusive events in Venn Diagram to show importance of pairwise disjoint.

- $P(\mathcal{S}) = 1$.

Venn Diagrams III

Venn diagrams for mutually exclusive.

Consequences of axioms

- $P(A) + P(A') = P(A \cup A') = P(\mathcal{S} = 1.$

- So often use $P(A) = 1 - P(A')$.

- $P(A \cup B) = P(A) + P(B) - P(AB)$.

- $P(A \cup B \cup C) = P(A) + P(B) + P(C) - P(AB) - P(AC) - P(BC) + P(ABC)$.

- If I did $A \cup B \cup C \cup D$ how many terms? What would last term be?

Where do probabilities come from?

- Counting for equally likely outcomes.
- Roll fair die. 6 sides. Each side has chance 1/6 (on one roll).
- Toss coin – 2 outcomes chance 1/2 each.
- Toss coin 3 times. 8 outcomes. All equally likely, chance 1/8.
- Same experiment. E is event first and third tosses heads.
- So

$$E = \{ \qquad\qquad\qquad \}.$$

- So $P(E) =$.

Long Run Relative Frequency

- Throw a pair of dice. Repeat – many, many times.
- 36 outcomes each time.
- Look for the event: sum is 10.
- 3 ways: 6,4; 5,5; 4,6. Chance is 3/36.
- Look at first n tosses.
- Number of 10s is *frequency* of "10" in first n tosses.
- *Relative frequency*: Frequency divided by n.

- Long run relative frequency is

$$P(10) = \frac{3}{36} = \lim_{n \to \infty} \frac{\text{number of 10s in first } n \text{ trials}}{n}.$$

- Both *definition* and *theorem*: Law of Large Numbers.

Subjective Probability, Bayesian

- Other views of probability.
- Subjective = "degree of belief".
- Apply probabilities to "events" like
  - Speed of sound less than 350 m/sec.
  - Mass of text more than 300 g.
  - Canucks win Stanley cup in 2015.
  - Richard lives to be 80.
- School of thought originated by Reverend Thomas Bayes.
- Probabilities are different for different people for same event.

Conditional Probability

- If A and B are events and $P(B) > 0$ then the conditional probability of A given B is
$$P(A|B) = \frac{P(AB)}{P(B)}.$$

- Idea is fraction of times where A occurs among times where B occurs.

- So
$$P(AB) = P(A|B)P(B)$$

- If $P(A) > 0$ also then
$$P(AB) = P(A|B)P(B) = P(B|A)P(A).$$

Independence

- Two events A and B are independent if

$$P(AB) = P(A)P(B).$$

- If $P(B) > 0$ and A and B are independent then

$$P(A|B) = \frac{P(AB)}{P(B)} = \frac{P(A)P(B)}{P(B)} = P(A).$$

- Knowing B happened didn't change chance that A happens.

- If $P(B) > 0$ and $P(A|B) = P(A)$ then

$$P(AB) = P(A|B)P(B) = P(A)P(B)$$

so A and B are independent.

- Three events A, B, and C are independent if

$$
\begin{aligned}
P(AB) &= P(A)P(B) \\
P(AC) &= P(A)P(C) \\
P(BC) &= P(B)P(C) \\
P(ABC) &= P(A)P(B)P(C)
\end{aligned}
$$

An example

- Toss a fair coin twice.
- Let A be the event that the first toss is heads.
- Let B be the event that the second toss is heads.
- Let C be the event that the two tosses are different

- Then $AB = \{HH\}$, $AC = \{HT\}$, $BC = \{TH\}$ and $ABC = \emptyset$.

- And $P(A) = P(B) = P(C) = 1/2$ so $P(AB) = P(A)P(B)$, $P(AC) = P(A)P(C)$ but $P(ABC) = 0 \neq P(A)P(B)P(C)$.

- So A and B are independent and so are A and C and so are B and C but A, B, and C are not independent.

Roles of independence and conditioning

- Even more formulas for 4 events, etc.

- Ideas are used in several ways.

- We often *model* the behaviour of a system by thinking about how A influences B.

- If there is no way for occurrence of A to affect occurrence of B and vice versa we *assume* A and B are independent.

- Sometimes we go the other way: we prove A and B are independent by checking the definition.

Conditional Independence, language

- We often make more subtle assumptions like "conditional independence':

$$P(A|BC) = P(A|B)$$

  means A and C are "conditionally independent" given B because

$$P(AC|B) = P(A|B)P(C|B).$$

- A word about language. We sometimes say "A is independent of B" instead of "A and B are independent". This sounds asymmetric but the idea is not.

- Sometimes we say "$A_1, \ldots, A_n$ are independent *of each other*" which means the same as "$A_1, \ldots, A_n$ are independent".

Combinatoric coefficients

- Toss coin n times: 2^n outcomes. All have same chance: $1/2^n = 2^{-n}$.

- Chance of (exactly) k heads?

- Must count number of outcomes with k heads.

- Pick k places from n to put H in seq of k Hs and $n-k$ Ts.

- So

$$P(\text{exactly } k \text{ heads in } n \text{ tosses}) = \binom{n}{k}\frac{1}{2^n}.$$

- Other notations

$$C_{n,k} \quad C_{k,n} \quad {}_nC_k = \binom{n}{k}.$$

- Number of permutations of n objects is $n!$.

- Number of ways to order k of n objects is

$$\frac{n!}{(n-k)!} = n \cdot (n-1) \cdots (n-k+1).$$

- n choices for first in list, $n-1$ for second, etc.

Example 1

One hard example. Shuffle 52 card poker deck. Deal 5. Find Chance of 3 of a kind.

Example 2

I shuffle a deck of cards, take one off the top, put it aside. What is chance next one is ace of spades?

Screening Programs

- Imagine serious condition (HIV +ve, drug use, prostate cancer).
- Population prevalence is 1%.
- Have imperfect test. (All real tests are imperfect.)
- Test has sensitivity 90%:

  $P(\text{positive test result}|\text{have condition}) = 0.9.$

- Test has very good specificity – 99%.

$$P(\text{neg test result}|\text{do not have condition}) = 0.99.$$

- Now suppose a random person takes the test and the result is positive. Does the person have the condition?

- Answer by computing

$$P(\text{have condition}|\text{positive test result})$$

Screening — numerical results

- Introduce notation. DON'T SKIP THIS STEP.

- Let B be the event "the selected person has the condition".

- Let A be the event "the test result is positive".

- We are given

$$
\begin{aligned}
P(B) &= \\
P(A|B) &= \\
P(A|B') &= 1 - P(A'|B') =
\end{aligned}
$$

- So $P(AB) =$.

- And $P(AB') =$

- So $P(A) =$

- We want $P(B|A)$.

Bayes' Theorem

- So

$$P(B|A) = \frac{P(BA)}{P(A)}$$
$$=$$

- Example of general strategy.

- Use Bayes' Theorem to switch order of conditions:

$$\begin{aligned} P(B|A) &= \frac{P(AB)}{P(A)} \\ &= \frac{P(A|B)P(B)}{P(AB) + P(AB')} \\ &= \frac{P(A|B)P(B)}{P(A|B)P(B) + P(A|B')P(B')} \end{aligned}$$

Tree Diagram Version

- To be done by drawing tree in class.
- At top of tree two possibilities: have condition or not.
- Left branch is "have condition" with probability 0.01.
- Right branch is "do not have condition with probability 0.99.
- Next layer has two branches at end of each top branch: left for test is positive, right for test is negative.
- Probabilities are, left to right, 0.9, 0.1, 0.01, 0.99

- Four paths to bottom:
  have/pos, have/neg, no/pos, no/neg
  with probs
  0.01*0.9, 0.01*0.1, 0.99*0.01, 0.99*0.99

- Nodes at bottom corresponding to test positive are first and third.

- Interested in probability of first branch given either first or third: