Final Review

Alexander Van Roijen

November 30, 2018

Please complete the following:

1. Lecture 1. Basics of Probability

   (a) **Definition (Sample space and event).** The sample space S of an experiment is the set of all possible outcomes of the experiment. An event A is a subset of the sample space S , and we say that A occurred if the actual outcome is in A.

   (b) **Definition (General definition of probability).** A probability space consists of a sample space S and a probability function P(.) which takes an event A $\subset$ S as input and returns P(A), a real number between 0 and 1, as output. The probability function must satisfy the following axioms: $P(\emptyset) = 0\, P(S) = 1$ and for a union of disjoint events, we get $P(A_1 \cup ...A_n) = P(A_1) + ...P(A_n)$

   (c) **Theorem. Properties of probability.** A probability function has the following properties, for any events A and B.

      i. $P(A^c) = 1 - P(A)$

      ii. if $A \subset BP(A) \leq P(B)$

      iii. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$ which can be extended $P(A \cup B \cup C) = P(A)+P(B)+P(C)-P(A \cap B)-P(A \cap C)-P(B \cap C)-P(A \cap B \cap C)$

   (d) **Definition (Conditional probability).** If A and B are events with $P(B) > 0$, then the conditional probability of A given B, denoted by $P(A|B)$, is defined as: $P(A|B) = \frac{P(A \cap B)}{P(B)}$ Further note that all probabilities are in fact conditional. We like to think of P(A) as our prior beliefs of an event, and P(A—B) as our posterior, or what we think it is given something is already known.

   (e) **Theorem.** For any events A1,..An with positive probabilities, $P(A1, ...An) = P(A1)P(A2|A1)P(A3|A1, A2)...P(An|A1, ...A_{n-1})$

   (f) **Theorem (Bayes rule).** $P(A|B) = \frac{P(B|A)P(A)}{P(B)}$

   (g) **Theorem (Law of total probability (LOTP)).** Let A1,...An be a partition of the sample space S with $P(Ai) > 0$ for all i. Then $P(B) = \sum_{i=1}^{n} P(B|A_i)P(A_i)$

   When we condition on an event E, we update our beliefs to be consistent with this knowledge, effectively putting ourselves in a universe where we know that E occurred. Within our new universe, however, the laws of probability operate just as before. Conditional probability satisfies all the properties of probability!

(h) **Theorem (Bayes rule with extra conditioning).** Provided that $P(A \cap E) > 0$ and $P(B \cap E) > 0$, we have $P(A|B,E) = \frac{P(B|A,E)P(A|E)}{P(B|E)}$

(i) T**heorem (Law of total probability (LOTP) with extra conditioning).** Let A1... An be a partition of the sample space S with $P(Ai|E) > 0$ for all i. Then $P(B|E) = \sum_{i=1}^{n} P(B|A_i,E)P(A_i|E)$

(j) **Definition (Independence of two events).** Events A and B are independent if P(A B) = P(A)P(B). If P(A) ¿ 0 and P(B) ¿ 0, then this is equivalent with P(A j B) = P(A), and also equivalent with P(B j A) = P(B). Independence is a symmetric relation.

(k) **Proposition.** If A and B are independent, then A^c and B are independent, A^c and B^c are independent, and A and B^c are independent.

(l) Definition (Independence of three events). Events A, B and C are said to be independent if all of the following relations hold:
$P(A \cap B) = P(A)P(B)$;
$P(A \cap C) = P(A)P(C)$;
$P(B \cap C) = P(B)P(C)$;
$P(A \cap B \cap C) = P(A)P(B)P(C)$

(m) **Definition (Conditional independence).** Events A and B are said to be conditionally independent given E if $P(A \cap B|E) = P(A|E)P(B|E)$.

(n) Problems shown: Monty Hall, and Positive test of conditionitis and bayes rule

2. (a) **Definition (Random variable).** Given an experiment with sample space S , a random variable is a function from the sample space S to the real numbers R. It is common, but not required, to denote random variables by capital letters. P(X=x) = P(X=X(s))

(b) Discrete PMFs are non negative, and sum to one over their support.

(c) **Definition (Bernoulli distribution).** An random variable X is said to have a Bernoulli distribution with parameter p if P(X = 1) = p and P(X = 0) = 1 ∼ p, where $0 < p < 1$. We write this as X ∼ Bern(p).

(d) **Theorem.** Let X ∼ $Bin(n, p)$, and $q = 1 - p$ (often taken to denote the failure of a Bernoulli trial). Then $n - X \sim Bin(n, q)$.

(e) **Theorem (Hypergeometric PMF).** Consider an urn with w white balls and b blacks balls. We draw n balls out of the urn at random without replacement such that all the $\binom{w+b}{n}$ samples are equally likely. Let X be the number of white balls in the sample. Then X is said to have the Hypergeometric distribution with parameters w, b and n: $X \sim HGeom(w, b, n)$. Then the PMF of X is
$P(X = k) = \frac{\binom{w}{k}\binom{b}{n-k}}{\binom{b+w}{n}}$

(f) **Theorem.** If $X \sim HGeom(w, b, n)$ and $Y \sim HGeom(n, w + b - n, w)$, then X and Y have the same distribution.

(g) **Theorem.** If X ∼ Bin(n,p), Y ∼ Bin(m,p), and X is independent of Y, then the conditional distribution of $X|X + Y = r \sim$ hgeom(n,m,r).

(h) **Theorem (Binomial as a limiting case of the Hypergeometric).** If X∼ HGeom(w, b, n) and N = w+b approaches infinity such that p = w/(w + b) remains fixed, then the PMF of X converges to the Bin(n, p) PMF.

(i) **Theorem (PMF of g(X))**. Let X be a discrete random variable and $gR- > R$. Then the support of g(X) is the set of all y such that g(x) = y for at least one x in the support of X, and the PMF of g(X) is $P(g(X) = Y) = \sum_{x:g(x)=y} P(X = x)$

(j) **Definition (Independence of two random variables).** Random variables X and Y are said to be independent if $P(X \leq x, Y \leq y) = P(X \leq x)P(Y \leq y)$ for all $x, y \in R$. In the discrete case, this is equivalent to the condition P(X = x, Y = y) = P(X = x)P(Y = y); for all x in the support of X and all y in the support of Y.

Happy holidays!