Fall 2020, Stat 510 : Homework 3

Moulinath Banerjee
University of Michigan

Independence of random variables: Random variables $X_1, X_2, \ldots, X_p$ are said to be independent if for any p (Borel) subsets of $\mathbb{R}$, say $A_1, A_2, \ldots, A_p$,

$$P(X_1 \in A_1, X_2 \in A_2, \ldots, X_p \in A_p) = P(X_1 \in A_1)\, P(X_2 \in A_2) \ldots \ldots P(X_p \in A_p)\,.$$

Thus, the notion of independence of random variables is a natural extension of the notion of independent events.

1. Let T be an exponential random variable with parameter β and let W be a random variable independent of T which assumes the value 1 with probability 2/3 and the value -1 with probability 1/3. Find the density of $X = WT$. (10 points)
   Hint: It would help to split up the event $\{X \leq x\}$ as the union of $\{X \leq x, W = 1\}$ and $\{X \leq x, W = -1\}$.

2. A random variable V is said to have a distribution symmetric about 0 if the distribution of V is the same as that of $-V$. Let V be a continuous random variable with continuous density function f.

   (a) (a)Show that V is distributed symmetrically about 0 if and only if $f(t) = f(-t)$, for every t. In other words f is an even function.

   (b) (b) Show that $E(X) = 0$.

   (c) (c) Let B be a Bernoulli random variable with probability of success θ where $0 < \theta < 1$. Find the density function of the random variable $\overline{V} = (2B - 1)V$. (5 + 5 + 10= 20 points)

3. (a) If P_1 and P_2 are independent Poisson random variables with parameters λ_1 and λ_2, then show that $P_1 + P_2$ is also Poisson with parameter $\lambda_1 + \lambda_2$. Recall that if W follows Poisson(θ), then the p.m.f. of W is,

   $$P(W = m) = \frac{e^{-\theta}\theta^m}{m!}$$

   Hint: Write $P(P_1 + P_2 = m)$ as $\sum_{i=0}^{m} P(P_1 = i, P_2 = m - i)$ and proceed. (5 points)

   (b) Use this result repeatedly to show that if $X_1, X_2, ..., X_n$ are independent Poisson random variables with parameters $\lambda_1, \lambda_2, ..., \lambda_n$ respectively, then $S_n \equiv X_1 + X_2 + ... + X_n$ follows Poisson with parameter $\lambda_1 + \lambda_2 + ... + \lambda_n$. (5 points)

(c) Let $\Lambda_n = \lambda_1 + \lambda_2 + \ldots + \lambda_n$ and assume that $\Lambda_n \to \infty$. Find the limit distribution of $\tilde{S}_n = (S_n - \Lambda_n)/\sqrt{\Lambda_n}$ as $n \to \infty$ using moment generating functions. (10 points)

4. Let Y denote the number of failures incurred before the r'th success (r fixed) in a sequence of Bernoulli trials with success probability p. Show by using mgfs that as $r \to \infty$ and $p \to 1$ such that $r(1-p) \to \lambda$ for some $\lambda > 0$,
$$P(Y = y) \to e^{-\lambda}\,\frac{\lambda^y}{y!}\,.$$

5. Consider a population of voters of size N numbered $1, 2, \ldots, N$ where each voter will vote either D or R. Let N_D be the number of D's and N_R the number of R's with $p_D = N_D/N$, the proportion of D's in the population being of interest. A sample of size $n \leq N$ is drawn from the population *without replacement*. We think of the individuals as being drawn one by one, and once a person is selected in the sample, they are not eligible to be drawn again: i.e. at every stage, we draw an individual uniformly at random from the total population minus the set of people already chosen in the sample. Thus, a particular sample outcome looks like $(m_1, m_2, \ldots, m_n)$ where the m_i's are all distinct positive integers between 1 and N. By the nature of the sampling each such ordered n tuple of integers is equally likely. Let $X_1, X_2, \ldots, X_n$ be the random variables defined by $X_j = 1$ if the j'th individual is D and 0 otherwise.

   (a) Consider an l-tuple of indices $1 \leq i_1 < i_2 < \ldots < i_l \leq n$, and let $(\epsilon_1, \epsilon_2, \ldots, \epsilon_n)$ be a fixed sequence of 1's and 0's. Compute $P(X_{i_1} = \epsilon_1, X_{i_2} = \epsilon_2, \ldots, X_{i_l} = \epsilon_l)$.
   Hint: Your answer should not depend on $i_1, i_2, \ldots i_l$. Among all possible permutations of size n from a population of N distinct individuals, in how many permutations are D's in the positions i_j such that ϵ_j and R's in the remaining positions? The probability then is simply a ratio. Only the sum of the ϵ_i's should enter the final expression for the desired probability.

   (b) Find $P(S_n = m)$ based on the calculation in (a) where S_n is the sum of the X_i's where $m \leq n$. Your answer should be a Hypergeometric probability.
   Hint: If $S_n = m$, then for some $1 \leq i_1 < i_2 < \ldots < i_m \leq n$, we have

$X_{i_1} = X_{i_2} = \ldots = X_{i_l} = 1$. This probability can be computed from (a). Now you need to consider how many distinct sequences of such i_j's there are.

(c) In the case $n = N$, find $E(S_n/n)$ and $\text{Var}(S_n/n)$. (10 + 10 + 5 = 25 points)

6. Let $X \sim f(x-\theta_0)$ (i.e. X has density $f(x-\theta_0)$ for some fixed number θ_0) where f is a symmetric, strictly unimodal and continuous density with mode at 0: i.e. $f(0) > f(x)$ for all $x \neq 0$ and $f(x) = f(-x)$ for all x. Find the interval that maximizes $P(X \in I_h)$ where I_h is an interval of length $2h$, for some $h > 0$.