Homework 1 (Stats 620, Winter 2017)

Due Thursday Jan 19, in class

1. **(a)** Let N denote a nonnegative integer-valued random variable. Show that

$$\mathbb{E}[N] = \sum_{k=1}^{\infty} \mathbb{P}\{N \geq k\} = \sum_{k=0}^{\infty} \mathbb{P}\{N > k\}.$$

(b) In general show that if X is nonnegative with distribution F, then

$$\begin{aligned} \mathbb{E}[X] &= \int_0^\infty \overline{F}(x)\,dx, \\ \mathbb{E}[X^n] &= \int_0^\infty n x^{n-1} \overline{F}(x)\,dx. \end{aligned}$$

Comment: These identities will be useful later in the course.

Solution:

(a) Let

$$I_k = \begin{cases} 1 & \text{if } N \geq k \\ 0 & \text{else} \end{cases} \tag{1}$$

and

$$J_k = \begin{cases} 1 & \text{if } N > k \\ 0 & \text{else} \end{cases}$$

Then, $N = \sum_{k=1}^{\infty} I_k = \sum_{k=0}^{\infty} J_k$. Taking expectations gives

$$\mathbb{E}(N) = \sum_{k=1}^{\infty} \mathbb{E}(I_k) = \sum_{k=0}^{\infty} \mathbb{E}(J_k)$$

Since $E[I_k] = P[N \geq k]$ and $E[J_k] = P[N > k]$, the result is shown.

(b)

$$\mathbb{E}(X) = \int_0^\infty x\,dF(x) = -\int_0^\infty x\,d\overline{F}(x)$$

Integration by parts gives

$$\mathbb{E}(X) = -x\overline{F}(x)\big|_0^\infty + \int_0^\infty \overline{F}(x)\,dx$$

Since $\mathbb{E}(X) < \infty$, $\lim_{x\to\infty} x\overline{F}(x) = 0$, giving

$$\mathbb{E}(X) = \int_0^\infty \overline{F}(x)\,dx$$

The second part of (b) is a very similar integration by parts.

2. Let X_n denote a binomial random variable, $X_n \sim \text{Binomial}\,(n, p_n)$ for $n \geq 1$. If $np_n \to \lambda$ as $n \to \infty$, show that

$$\mathbb{P}\{X_n = i\} \to e^{-\lambda}\lambda^i/i! \qquad \text{as } n \to \infty.$$

Hint: Write out the required binomial probability, expanding the binomial coefficient into a ratio of products. Taking logarithms may be helpful to show that $\lim_{n\to\infty} c_n = c$ implies $\lim_{n\to\infty}(1 - c_n/n)^n = e^{-c}$.

Solution:

$$\begin{aligned}
\mathbb{P}(X_n = i) &= \binom{n}{i} {p_n}^i (1-p_n)^{n-i} \\
&= \frac{n \times n-1 \times \cdots \times n-i+1}{i!} \times \frac{(np_n)^i}{n^i} \times \frac{(1-p_n)^n}{(1-p_n)^i}
\end{aligned}$$

As $n \to \infty$, $(n-k)/n \to 1$ for $k = 0, 1, \cdots, i-1$ and so $\lim_{n\to\infty} n!/(n-i)!n^i = 1$. Define $\lambda_n = np_n$. Now, $np_n \to \lambda$ implies $(np_n)^i \to \lambda^i$ and $(1-p_n)^i \to 1$ for all fixed i. Now,

$$\begin{aligned}
\log \lim_{n\to\infty} (1 - \lambda_n/n)^n &= \lim_{n\to\infty} n \log(1 - \lambda_n/n) \\
&= \lim_{n\to\infty} n(-\lambda_n/n + o(1/n)) \qquad (2) \\
&= -\lim_{n\to\infty} \lambda_n = -\lambda
\end{aligned}$$

So, $\lim_{n\to\infty}(1 - \lambda_n/n)^n = \exp(-\lambda)$,where (2) is justified by a Taylor expansion of $\log(1-x)$ and $a_n = o(b_n)$ means $a_n/b_n \to 0$.

3. Let F be a continuous distribution function and let U be a uniformly distributed random variable, $U \sim \text{Uniform}\,(0, 1)$.

(a) If $X = F^{-1}(U)$, show that X has distribution function F.

(b) Show that $-\log(U)$ is an exponential random variable with mean 1.

Comment: Part (b) gives a way to simulate exponential random variables using a computer with a random number generator producing $U[0, 1]$ random variables.

Solution:

(a)

$$\begin{aligned}
\mathbb{P}(X \leq x) &= \mathbb{P}(F^{-1}(U) \leq x) \\
&= \mathbb{P}(U \leq F(x)) \qquad (3) \\
&= \int_0^{F(x)} du = F(x)\,.
\end{aligned}$$

Note that in (3), we used the fact that F is non-decreasing. Otherwise the equality does not necessarily hold. For instance let $F(t) = \exp\{-t\}$ and note that $F(0) = 1$, $\lim_{t\to\infty} F(t) = 0$ and F is decreasing. In this case, $\mathbb{P}(F^{-1}(U) \leq x) = \mathbb{P}(U \geq F(x))$.

(b) Note that U has the same distribution as $1 - U$. If $F(x) = 1 - e^{-x}$, then $F^{-1}(x) = -\log(1-x)$ so (a) shows that $-\log(U)$ has c.d.f. e^{-x}, i.e. X is exponentially distributed with parameter $\lambda = 1$. Thus, $\mathbb{E}(X) = 1/\lambda = 1$.

4. Let $f(x)$ and $g(x)$ be probability density functions, and suppose that for some constant c, $f(x) \leq cg(x)$ for all x. Suppose we can generate random variables having density function g, and consider the following algorithm.

   Step 1. Generate Y, a random variable having density function g.

   Step 2. Generate $U \sim$ Uniform $(0, 1)$.

   Step 3. If $U \leq \frac{f(Y)}{cg(Y)}$ set $X = Y$. Otherwise, go back to Step 1.

   Assuming that successively generated random variables are independent, show that:

   (a) X has density function f.

   (b) the number of iterations of the algorithm needed to generate X is a geometric random variable with mean c.

   Comment: The procedure investigated in this problem is a standard computational tool for simulating a random variable with a given target density $f(x)$.

   Solution: **(b)** Define infinite sequences $\{Y_n, n \geq 1\}$ and $\{U_n, n \geq 1\}$ where Y_n are i.i.d. with density g and U_n are iid uniform $(0, 1)$. Define $E_n = \{U_n \leq f(Y_n)/cg(Y_n)\}$. E_n is a sequence of independent trials and the algorithm stops at the first "success." Thus the number of iterations is geometric with parameter

$$\begin{aligned} p &= \mathbb{P}(E_n) = \mathbb{E}(\mathbb{P}(E_n|Y_n)) \\ &= \int_{-\infty}^{\infty} \frac{f(y)}{cg(y)} g(y)\, dy = \frac{1}{c}. \end{aligned}$$

   The mean number of iterations is $1/p = c$.

   (a)

$$\begin{aligned} \mathbb{P}(Y_n \leq x|E_n) &= \mathbb{P}(Y_n \leq x, E_n)/\mathbb{P}(E_n) \\ &= c\,\mathbb{E}(\mathbb{P}(Y_n \leq x, E_n|Y_n)) \\ &= c\int_{-\infty}^{x} \frac{f(y)}{cg(y)} g(y)\, dy = F(x), \end{aligned}$$

   where $F(x) = \int_{-\infty}^{x} f(u)\, du$. On the nth iteration, conditional on the algorithm terminating at that iteration, Y_n has conditional c.d.f. F, so X has density f.

5. If $X_1, X_2, \ldots, X_n$ are independent and identically distributed exponential random variables with parameter λ, show that $S = \sum_{i=1}^{n} X_n$ has a gamma distribution with parameters (n, λ). That is, show that the density function of S is given by

$$f(t) = \lambda e^{-\lambda t}(\lambda t)^{n-1}/(n-1)!, \qquad t \geq 0.$$

   Instruction: Use moment generating functions for this question.

   Solution: Let $Y = \sum_{i=1}^{n} X_i$. Since $X_i \sim$ Exponential (λ), the MGF of X_i is

$$\mathbb{E}(e^{tX_i}) = \lambda/(\lambda - t)\,, 0 \leq t < \lambda\,.$$

Then

$$\begin{aligned}
\mathbb{E}(e^{tY}) &= \mathbb{E}(\exp\{t\sum_{i=1}^{n} X_i\}) \\
&= \mathbb{E}(\prod_{i=1}^{n} e^{tX_i}) \\
&= \prod_{i=1}^{n} \mathbb{E}(e^{tX_i}) \qquad \text{using independence} \\
&= \lambda^n/(\lambda - t)^n \qquad \text{for } 0 \leq t < \lambda.
\end{aligned}$$

This can be recognized as the MGF of a Gamma (n, λ) random variable. Since the MGF, when exists, determines a distribution uniquely, the result is proved.

Recommended reading:
These homework problems derive from Chapter 1 of Ross "Stochastic Processes," all of which is relevant material. There are too many examples to study them all! Some suggested examples are 1.3(A), 1.3(C), 1.5(A), 1.5(D), 1.9(A).

Supplementary exercises: 1.5, 1.22.
These are optional, but recommended. Do not turn in solutions—they are in the back of the book. To prove Booles's inequality (B5 in the notes for the first class), one can write $\bigcup_{i=1}^{\infty} E_i$ as a disjoint union via defining $F_i = E_i \cap E_{i-1}^c \cap \cdots \cap E_1^c$.