Assignment #8

Due on Monday, October 31, 2016

Read Section 4.1.5 on *The Poisson Distribution* in the class lecture notes at `http://pages.pomona.edu/~ajr04747/`.

Read Section 4.1.6 on *Estimating Mutation Rates in Bacterial Populations* in the class lecture notes at `http://pages.pomona.edu/~ajr04747/`.

Read Section 4.2 on *Random Processes* in the class lecture notes at `http://pages.pomona.edu/~ajr04747/`.

Do the following problems.

1. Assume that Y is a Poisson random variable with mean $\lambda > 0$. Compute the variance of Y, $\text{Var}(Y) = E(Y^2) - [E(Y)]^2$.

2. Let $M(t)$ denote number of bacteria in a colony of initial size N_o that develop a certain type of mutation in the time interval $[0, t]$. It was shown in the lectures that if there are no mutations at time $t = 0$, and if $M(t)$ follows the assumptions of a Poisson process, then the probability of no mutations in the time interval $[0, t]$ is given by
$$P_0(t) = P[M(t) = 0] = e^{-\lambda t}$$
where $\lambda > 0$ is the average number of mutations per unit time.

   Let $T > 0$ denote the time at which the first mutation occurs.

   (a) Explain why T is a random variable. Observe that it is a *continuous* random variable.

   (b) For any $t > 0$, explain why the statement
$$P[T > t] = P[M(t) = 0]$$
is true, and use it to compute
$$F(t) = P[T \leq t].$$
The function $F(t)$, usually denoted by $F_T(t)$, is called the *cumulative distribution function,* or cdf, of the random variable T.

   (c) Compute the derivative $f(t) = F'(t)$ of the cdf F obtained in the previous part.

   The function $f(t)$, usually denoted by $f_T(t)$, is called the *probability density function,* or pdf, of the random variable T.

3. Given a continuous random variable X with pdf f_X, the *expected value* of X is defined to be
$$E(X) = \int_{-\infty}^{\infty} x f_X(x) dx.$$
Use this formula to compute the expected value of the T, where T is the random variable defined in Problem 2; that is, $T > 0$ is he time at which the first mutation occurs for a bacterial colony exposed to a virus at time $t = 0$, assuming that there are no mutations at that time. How does this value relate to the average mutation rate λ?

4. *Modeling Survival Time after a Treatment.* Consider a group of people who have received a treatment for a disease such as cancer. Let T denote the *survival time;* that is, T is the number of years a person lives after receiving the treatment.

   Assume that the probability that a person receiving the treatment at time t will not survive past time $t + \Delta t$ is proportional to Δt; denote the constant of proportionality by $\mu > 0$. If we let $p(t)$ denote the probability that a person who received the treatment at time $t_o = 0$ is still alive at time t, obtain a differential equation for $p(t)$ and solve for $p(t)$ assuming that $p(0) = 1$.

5. *Modeling Survival Time after a Treatment, Continued.* Let T, μ and $p(t)$ be as in Problem 4.

   (a) Explain why
   $$\Pr(T > t) = p(t).$$

   (b) Give a formula for computing
   $$F_T(t) = \Pr(T \leqslant t), \qquad \text{for all } t > 0.$$
   $F_T(t)$, is called the *cumulative distribution function,* or cdf, of the random variable T.

   (c) Let $f_T(t) = F_T'(t)$ for all $t > 0$. Show that f_T is of the form
   $$f_T(t) = \begin{cases} \dfrac{1}{\beta} e^{-t/\beta} & \text{for } t \geqslant 0 \\ 0 & \text{for } t < 0, \end{cases}$$
   for some positive constant β.
   What is β in terms of μ?

   (d) Find the expected value of T; that is, compute $E(T) = \displaystyle\int_{-\infty}^{\infty} t f_T(t)\ \mathrm{d}t$.