ENEE 324

Sections 0101, 0102, FR01 Spring 2011

Homework #9

Remember, the goal here is to understand ***how*** *to do the calculations, not to produce the correct answer using pure instinct. Please explain* ***how*** *you have arrive at your answers.*

Please attach all your Matlab code.

1) Let X be a random variable with the following CDF: $F_X(x) = \begin{cases} 0 & x < 0 \\ x^3/64 & 0 \le x < 4 \\ 1 & x \ge 4 \end{cases}$

   a) Find and sketch the PDF $f_X(x)$.

   b) What is $P[X > 2]$?

   c) Derive an expression for the conditional PDF $f_{X|B}(x)$, where B is the event that $X > 2$.

   d) Calculate $E[X|B]$ and $VAR[X|B]$.

1) In homework 7, you were told the following: John Cook, while exploring a remote tropical island encounters a very large chest filled with gold. It is so large in fact, that it will require three persons to carry it back to the ship (him and two others). Suppose that sailors arrive at the treasure site randomly in a Poisson manner with an average rate of $\lambda = 1$ sailors per hour. Let T be a continuous random variable describing the time that John Cook must wait until he has enough people to move the treasure.

   a) Suppose John Cook has been waiting for an hour and no sailor has yet arrived. What is the conditional PDF of T, given this information?

   b) Similarly, given that he has already waited an hour and no sailor has yet arrived, how much longer, on average, should he expect to wait for a enough sailors to move the chest?

2) The function $\Phi(z)$ is the CDF for a standard normal random variable. Many numerical packages, such as basic Matlab, do not provide it but instead provide a similar function called the error function,

   $\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-u^2)\,du$; its relationship to $\Phi(z)$ was shown in class.

   a) Find an expression for $\Phi^{-1}(p)$ in terms of the inverse error function $\mathrm{erf}^{-1}(x)$.

   b) A company manufactures a gasket with a nominal size of 1 cm. The manufacturing process is not perfect, and the actual size X is modeled as a Gaussian random variable with a mean of 1 cm and standard deviation σ. If the size of the gasket is off by more than 0.5 mm, it will not function. The manufacturer hires you to answer this: How small must σ be in order to insure that the likelihood of the part malfunctioning is less than 1/10000? Use the standard Matlab function `erfinv()` for $\mathrm{erf}^{-1}(x)$.

3) A Pareto distribution is useful for describing random variables where a large fraction of a resource (e.g. money or time) is dominated by a small fraction of the population. This may be informally described as obeying the "80-20" rule, e.g. 80% of your time is spend on 20% of a problem set. Formally it is described by a normalized inverse power law (that vanishes below some minimum value $x_{\min}$). So

$X \sim$ Pareto($\alpha, x_{\min}$) means that $f_X(x) = \dfrac{\alpha}{x_{\min}}\left(\dfrac{x_{\min}}{x}\right)^{\alpha+1} u(x - x_{\min})$, where $u()$ is the unit step function.

It can also be shown that $E[X] = \dfrac{\alpha x_{\min}}{\alpha - 1}$ and $VAR[X] = \dfrac{\alpha x_{\min}^2}{(\alpha-1)^2(\alpha-2)}$.

a) Compute the CDF, $F_X(x)$, analytically (i.e. without Matlab). Because this may be trickier than it seems, verify at each step that the units are correct (that the PDF has units of a probability density, $\propto x^{-1}$ or $\propto x_{\min}^{-1}$, and that the CDF has units of pure probability). Hint: Use a variable substitution of $u = x/x_{\min}$ while being careful to apply the variable substitution to the integral limits also.

b) Compute the inverse CDF, $F_X^{-1}(p)$, analytically (i.e. without Matlab).

c) For the case $x_{\min} = 1$ (which is equivalent to measuring x in units of $x_{\min}$) and $\alpha = 6$, use Matlab to plot the PDF for the range [0,5] in steps of 0.01. In the plot, set the x-axis limits to range from 0 to 5, and the y-axis limits to range from 0 to α, for comparison below.

d) Using Matlab and the inverse CDF computed above, generate 100000 realizations of this Pareto(6,1) random variable. Plot the histogram using bins of the form `[0:0.01:10]` (set the x-axis limits to range from 0 to 5, but set the y-axis limits to range from 0 to 1000α). How closely does the shape of the histogram match the shape of the PDF? [If you superimpose the PDF (scaled up by a factor of 1000) on the histogram (using `hold on` and `hold off`) the two should match cleanly.]

e) Using Matlab, what is the mean of these realizations? How does it compare to the theoretical mean?

f) Similarly, what is the variance of these realizations? How does it compare to the theoretical variance?

g) Using Matlab, estimate the expectation value of the square root of this random variable.

4) Suppose X is an exponential random variable with a mean of 1. Y is another random variable that, given X, has an exponential distribution with a mean of $1/X$: $f_{Y|X}(y \mid x) = x\,\mathrm{e}^{-xy}\,u(y)$.

a) Find an expression for the joint PDF $f_{X,Y}(x,y)$.

b) Find an expression for the marginal PDF $f_Y(y)$. Hint: you may need: $\int_0^\infty u\,\mathrm{e}^{-au}\,du = a^{-2}$.