Calculus Challenge #10

Solutions due March 16, 2011

Scientists and engineers are often interested in the results of random processes. The time between phone calls at a Help Desk, or the lifespan of a component in a machine are classic examples. These quantities are called random variables and the probability that a random variable takes on a certain value is given by a probability density function, or pdf. For example, the probability that the next phone call to the Help Desk arrives between 2 and 5 minutes from now is given by the area under the pdf from 3 to 5, or

$$P\left(3 \le x \le 5\right) = \int_3^5 pdf\,(x)\,dx\,.$$

The probabilities are given by the area under the pdf over the interval of interest. This is one clear case where area under a curve has real meaning.

A previous problem using pdf's for the Normal Distribution and the *t*-distribution is Problem # 6 from 2008-2009 (see http://courses.ncssm.edu/math/POW/POW08_09/Calculus%20Challenge%20Problem%206.pdf) .

A probability density function, *f*, must satisfy two specific criterion:

1) $f\left(x\right) \ge 0$ for all *x*.

2) $\int_{-\infty}^{\infty} f\,(x)\,dx = 1.$

One of the most important pdf's is the exponential distribution. The time between calls at a Help Desk can be modeled with the exponential pdf

$$f\left(x\right) = \begin{cases} \lambda e^{-\lambda x} \text{ if } x \ge 0 \\ 0 \text{ if } x < 0 \end{cases}.$$

The parameter λ is called the *intensity* and its value depends upon how busy or active the phone lines at the Help Desk are. For example, if $\lambda = 3$, the probability that the next call will arrive within the next two minutes is $P\left(0 \le x \le 2\right) = \int_0^2 3e^{-3x}\,dx = 0.997$ (almost certainly, a call will come within two minutes), while if $\lambda = 0.15$, the probability that the next call will arrive within the next two minutes is $P\left(0 \le x \le 2\right) = \int_0^2 0.15e^{-0.15x}\,dx = 0.259$ (74% of the time, you will wait at least two minutes for the next call). We will generate an interpretation for the value of λ as a part of this challenge.

1) Prove that $f\left(x\right) = \begin{cases} \lambda e^{-\lambda x} \text{ if } x \ge 0 \\ 0 \text{ if } x < 0 \end{cases}$ is a legitimate pdf satisfying the two criterion listed above.

2) The average value of the random variable, denoted μ, is defined as $\mu = \int_{-\infty}^{\infty} x \cdot f(x)\,dx$. The variance, denoted σ^2, is defined as $\sigma^2 = \int_{-\infty}^{\infty} (x-\mu)^2 \cdot f(x)\,dx$. Find μ (the average time between calls) and σ^2 for the exponential distribution $f(x) = \begin{cases} \lambda e^{-\lambda x} \text{ if } x \geq 0 \\ 0 \text{ if } x < 0 \end{cases}$. The standard deviation σ is the square root of the variance. What does the value of λ tell you about the process being modeled?

3) The median, *m*, is the value of the random variable for which the area under the pdf to the left of *m* and to the right of *m* are each one-half. The inter-quartile range is the length of the interval between the value x_1 with $P(x \leq x_1) = 0.25$ and the value x_2 with $P(x \geq x_2) = 0.25$. Find the median and IQR for the exponential distribution $f(x) = \begin{cases} \lambda e^{-\lambda x} \text{ if } x \geq 0 \\ 0 \text{ if } x < 0 \end{cases}$.

4) If the average time between calls at the Help Desk is 2 minutes, what is the probability that you can take a 1 minute break and not miss a call?

5) The probability that the next call will come in the next minute is given by $\int_0^1 \lambda e^{-\lambda x}\,dx$. The probability that the next call will arrive in the 5th minute from now is given by $\int_5^6 \lambda e^{-\lambda x}\,dx$. The conditional probability that the next call will arrive in the 5th minute given that no call has arrived in the first five minutes is given by $P(5 \leq x \leq 6 \mid x \geq 5) = \dfrac{\int_5^6 \lambda e^{-\lambda x}\,dx}{\int_5^{\infty} \lambda e^{-\lambda x}\,dx}$. Determine these three probabilities.

6) The exponential distribution has what is known as the "memoryless" property. Show that the probability of a call arriving in the next *k* minutes is the same as the probability that a call arrives in *k* additional minutes given that no call has arrived in the first *n* minutes. Explain why this is called "memoryless".