Calculus Challenge #10 ***SOLUTIONS***

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

1) $f(x) \geq 0$ 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

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

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 $P(0 \leq x \leq 2) = \int_0^2 3e^{-3x}\,dx = 0.997$ (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 $P(0 \leq x \leq 2) = \int_0^2 0.15e^{-0.15x}\,dx = 0.259$ (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 $f(x) = \begin{cases} \lambda e^{-\lambda x} \text{ if } x \geq 0 \\ 0 \text{ if } x < 0 \end{cases}$ is a legitimate pdf satisfying the two criterion listed above.

Clearly, since f is an exponential, its values are positive for $x \geq 0$ and zero elsewhere, so condition 1 is satisfied. For the second condition, we evaluate the improper integral $\int_{-\infty}^{\infty} f(x)dx = 0 + \int_0^{\infty} \lambda e^{-\lambda x}dx = -\left(\frac{\lambda}{\lambda}\right)e^{-\lambda x}\Big|_0^{x=\infty} = \lim_{k\to\infty}\left(1 - e^{-\lambda k}\right) = 1.$

2) The average value of the random variable, denoted μ, is defined as $\mu = \int_{-\infty}^{\infty} x \cdot f(x)\,dx$.

This requires integration by parts, so $\mu = \int_0^\infty x \cdot e^{-\lambda x}\, dx = -xe^{-\lambda x}\Big|_0^\infty + \int_0^\infty e^{-\lambda x}\, dx.$ The first term goes to zero, so we have only $\int_0^\infty e^{-\lambda x}\, dx = \frac{1}{\lambda}.$ The mean is the reciprocal of lambda. So, if the average time between calls is 1.5 minutes, then the probability of a call in the next two minutes would be $\int_0^2 \left(\frac{2}{3}\right) e^{-\left(\frac{2}{3}\right)x}\, dx.$

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?

Again, we use integration by parts, only this time we need to do it more times.

$\sigma^2 = \int_0^\infty \left(x - \frac{1}{\lambda}\right)^2 \cdot \lambda e^{-\lambda x}\, dx$. Let $u = \left(x - \frac{1}{\lambda}\right)^2$ and $dv = \lambda e^{-\lambda x} dx$, so

$\int_0^\infty \left(x - \frac{1}{\lambda}\right)^2 \cdot \lambda e^{-\lambda x}\, dx = -\left(x - \frac{1}{\lambda}\right)^2 e^{-\lambda x}\Big|_0^\infty + 2\int_0^\infty \left(x - \frac{1}{\lambda}\right) e^{-\lambda x}\, dx$. The first term simplifies to just $\frac{1}{\lambda^2}$. The remaining integral can be simplified as $2\left(\int_0^\infty \left(xe^{-\lambda x}\right) dx - \frac{1}{\lambda}\int_0^\infty \left(e^{-\lambda x}\right) dx\right)$. We have already evaluated the first integral term there, $\int_0^\infty \left(xe^{-\lambda x}\right) dx = \left(\frac{1}{\lambda}\right)\int_0^\infty \left(x\lambda e^{-\lambda x}\right) dx = \left(\frac{1}{\lambda^2}\right)$, so this adds $\frac{2}{\lambda^2}$ to the first term. Finally, we have $-2\left(\frac{1}{\lambda}\int_0^\infty \left(e^{-\lambda x}\right) dx\right)$. This is $-\frac{2}{\lambda}\left(\frac{1}{\lambda}\right) = -\frac{2}{\lambda^2}$.

So, the value of $\sigma^2 = \int_0^\infty \left(x - \frac{1}{\lambda}\right)^2 \cdot \lambda e^{-\lambda x}\, dx$ is $\frac{1}{\lambda^2} + \frac{2}{\lambda^2} - \frac{2}{\lambda^2} = \frac{1}{\lambda^2}$. The variance of the exponential distribution is $\frac{1}{\lambda^2}$ and the standard deviation is $\frac{1}{\lambda}$.

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

The median is fairly easy to find. Solve the equation $\int_0^m \lambda e^{-\lambda x}\,dx = \frac{1}{2}$ for m. So, $1 - e^{-m\lambda} = \frac{1}{2}$ and $e^{-m\lambda} = \frac{1}{2}$.

Then $m = \frac{\ln(2)}{\lambda}$.

Similarly, to find the quartiles, we have $\int_0^q \lambda e^{-\lambda x}\,dx = \frac{1}{4}$ and $\int_0^Q \lambda e^{-\lambda x}\,dx = \frac{3}{4}$, so $q = \frac{\ln(4) - \ln(3)}{\lambda}$ and $Q = \frac{\ln(4)}{\lambda}$. The IQR is $Q - q = \frac{\ln(3)}{\lambda}$.

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?

On a slow night, when the average time between calls is 2 minutes, the probability the next call comes in the next minute is $\int_0^1 \left(\frac{1}{2}\right) e^{-\left(\frac{1}{2}\right)x}\,dx = 0.3935$, so the probability that a call will not come is 0.6065.

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 $P(5 \le x \le 6 \mid x \ge 5) = \frac{\int_5^6 \lambda e^{-\lambda x}\,dx}{\int_5^\infty \lambda e^{-\lambda x}\,dx}$. Determine these three probabilities.

We have $\int_0^1 \lambda e^{-\lambda x}\,dx = 1 - e^{-\lambda}$, and $\int_5^6 \lambda e^{-\lambda x}\,dx = e^{-5\lambda} - e^{-6\lambda}$.

$P(5 \le x \le 6 \mid x \ge 5) = \frac{\int_5^6 \lambda e^{-\lambda x}\,dx}{\int_5^\infty \lambda e^{-\lambda x}\,dx} = \frac{e^{-5\lambda} - e^{-6\lambda}}{e^{-5\lambda} - 0} = 1 - e^{-\lambda}$. Notice that this is the same as in the next minute.

Every next minute acts just like the first minute.

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".

We know the probability of an arrival in the first *k* minutes is $P\left(0 \le x \le k\right) = \int_0^k \lambda e^{-\lambda x}\, dx = 1 - e^{-k\lambda}$. Also, we just learned that the probability of a first arrival in the interval (*k*, *k*+*n*) given that there was no arrival in the first *n* minutes is

$$P\left(n \le x \le n+k \mid x \ge n\right) = \frac{\int_n^{n+k} \lambda e^{-\lambda x}\, dx}{\int_n^{\infty} \lambda e^{-\lambda x}\, dx} = \frac{e^{-n\lambda} - e^{-(n+k)\lambda}}{e^{-n\lambda}} = 1 - e^{-k\lambda}$$

. These are the same probabilities.

As we saw in 5), if we have made it to n minutes without a call, it is just like we are at the beginning again. Every new minute is the first minute, hence "memoryless".