Homework 3

Due: Friday, Oct. 16, 2009

1. Let the density for X be given by

$$f(x) = ce^{-|x|} \qquad -\infty < x < \infty$$

(a) Find the value of c that makes this a density.

(b) Show that

$$\frac{1}{2}\int_{-\infty}^{\infty} |x|e^{-|x|}dx$$

exists.

(c) Find $E[X]$.

(d) Show that

$$M_X(s) = \frac{-1}{s^2-1} \qquad -1 < s < 1$$

(e) Use $M_X(s)$ to find $E[X]$ and $E[X^2]$.

(f) Find $Var[X]$.

Solution: (a) We must find c such that $\int_{-\infty}^{\infty} ce^{-|x|}dx = 1$

$$\begin{aligned} c\int_{-\infty}^{\infty} e^{-|x|}dx &= c\int_{-\infty}^{0} e^{x}dx + c\int_{0}^{\infty} e^{-x}dx \\ &= c(e^x|_{-\infty}^{0} + (-e^{-x})|_0^{\infty}) \\ &= c(1+1) = 2c \end{aligned}$$

$2c = 1 \Rightarrow c = 1/2$

(b)

$$\begin{aligned} &\frac{1}{2}\int_{-\infty}^{\infty} |x|e^{-|x|}dx = \frac{1}{2}\int_{-\infty}^{0}(-x)e^{x}dx + \frac{1}{2}\int_{0}^{\infty} xe^{-x}dx \\ &= \left(-\frac{1}{2}\right)\left(xe^x|_{-\infty}^{0} - \int_{-\infty}^{0} e^x dx\right) + \frac{1}{2}\left((-xe^{-x})|_0^{\infty} + \int_0^{\infty} e^{-x}dx\right) \\ &= \frac{1}{2} + \frac{1}{2} = 1 \end{aligned}$$

So, the specified integral exists.

(c)

$$\begin{aligned} E[X] &= \int_{-\infty}^{\infty} xf(x)dx = \int_{-\infty}^{\infty} x\left(\frac{1}{2}\right)e^{-|x|}dx \\ &= \frac{1}{2}\left(\int_{-\infty}^{0} xe^x dx + \int_0^{\infty} xe^{-x}dx\right) = 0 \end{aligned}$$

(d)

$$\begin{aligned}
M_X(s) = E[e^{sX}] &= \int_{-\infty}^{\infty} e^{sx}\left(\frac{1}{2}\right)e^{-|x|}dx \\
&= \frac{1}{2}\left(\int_{-\infty}^{0} e^{sx}e^{x}dx + \int_{0}^{\infty} e^{sx}e^{-x}dx\right) \\
&= \frac{1}{2}\left(\frac{1}{s+1}e^{(s+1)x}\Big|_{-\infty}^{0} + \left(-\frac{1}{(1-s)}\right)e^{-(1-s)x}\Big|_{0}^{\infty}\right) \\
&= \frac{1}{2}\left(\frac{1}{s+1} + \frac{1}{1-s}\right) = \frac{1}{1-s^2}
\end{aligned}$$

if $s+1>0$ and $1-s>0$, i.e. $-1<s<1$.

(e)

$$\begin{aligned}
E[X] = \frac{d}{ds}M_X(s)\Big|_{s=0} &= \frac{2s}{(1-s^2)^2}\Big|_{s=0} = 0 \\
E[X^2] = \frac{d^2}{ds^2}M_X(s)\Big|_{s=0} &= \frac{2(1-s^2)^2 - 8s^2(1-s^2)}{(1-s^2)^4}\Big|_{s=0} = 2
\end{aligned}$$

(f)

$$Var[X] = E[X^2] - (E[X])^2 = 2 - 0^2 = 2$$

2. Show that for $k>0$ and $\beta>0$,

$$\int_0^{\infty} f(x)dx = \int_0^{\infty} \frac{1}{\beta^k\Gamma(k)}x^{k-1}e^{-x/\beta}dx = 1$$

thereby showing that the function $f(x)$ is a density for a continuous random variable.

Proof: Let $z = \frac{x}{\beta}$. Then $x = \beta z$ and $dx = \beta dz$. Substitution yields

$$\begin{aligned}
&\int_0^{\infty} \frac{1}{\beta^k\Gamma(k)}x^{k-1}e^{-x/\beta}dx \\
&= \frac{1}{\beta^k\Gamma(k)}\int_0^{\infty}(\beta z)^{k-1}e^{-z}\beta dz \\
&= \frac{1}{\beta^k\Gamma(k)}\beta^k\int_0^{\infty} z^{k-1}e^{-z}dz \\
&= \frac{1}{\beta^k\Gamma(k)}\beta^k\Gamma(k) = 1
\end{aligned}$$

3. Among diabetics, the fasting blood glucose level X may be assumed to be approximately normally distributed with mean 106 milligrams per 100 milliliters and standard deviation 8 milligrams per 100 milliliters.

(a) Sketch a graph of the density for X. Indicate on this graph the probability that a randomly selected diabetic will have a blood glucose level between 90 and 122 mg / 100 ml. Find this probability.

(b) Find $P[X \leq 120\ mg/\ 100\ ml]$.

(c) Find the point that has the property that 25% of all diabetics have a fasting glucose level of this value or lower.

(d) If a randomly selected diabetic is found to have fasting blood glucose level in excess of 130, do you think there is cause for concern? Explain, based on the probability of this occurring naturally.

Solution: (a) The graph of the density for X is shown in Fig. 1

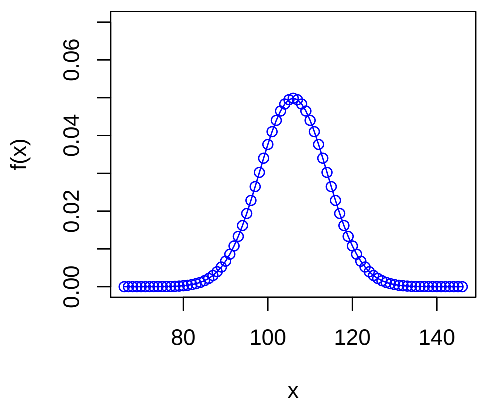

Figure 1: Problem 3 normal density

The required probability is

$$P[90 < X < 122] = F(122) - F(90) = .9545$$

(b)

$$P[X \leq 120] = F(120) = .9599$$

(c) The required point is X_0 such that $P[X \leq x_0] = .25$, i.e.the 25% quantile. X_0 can be obtained directly in many statistical package. Sample R code is attached. $x_0 = 100.60$.

(d) $P[X \geq 130] = 1 - P[X \leq 130] = 1 - F(130) = 0.0013$ which indicates that a fasting blood glucose level greater than 130 is quite abnormal.

4. A random variable X has the density function $f(x) = c/(x^2+1)$, where $-\infty < x < \infty$.

(a) Find the value of the constant c.

(b) Find the probability that X^2 lies between 1/3 and 1.

(c) Find the distribution function corresponding to the density function.

Solution: (a)

$$\begin{aligned}\int_{-\infty}^{\infty} \frac{c}{x^2+1}dx &= c\tan^{-1}x\big|_{-\infty}^{\infty} \\ &= c\left(\frac{\pi}{2}+\frac{\pi}{2}\right) = 1 \\ &\Rightarrow c = 1/\pi\end{aligned}$$

(b) If $\frac{1}{3} \leq X^2 \leq 1$, then either $\frac{\sqrt{3}}{3} \leq X \leq 1$ or $-1 \leq X \leq -\frac{\sqrt{3}}{3}$. Thus the required probability is

$$\begin{aligned}&\frac{1}{\pi}\int_{-1}^{-\sqrt{3}/3} \frac{1}{x^2+1}dx + \frac{1}{\pi}\int_{\sqrt{3}/3}^{1} \frac{1}{x^2+1}dx \\ &= \frac{2}{\pi}\int_{\sqrt{3}/3}^{1} \frac{1}{x^2+1}dx \\ &= \frac{2}{\pi}\left[\tan^{-1}(1) - \tan^{-1}\left(\frac{\sqrt{3}}{3}\right)\right] \\ &= \frac{2}{\pi}\left(\frac{\pi}{4}-\frac{\pi}{6}\right) = \frac{1}{6}\end{aligned}$$

(c)

$$\begin{aligned}F(x) &= \int_{-\infty}^{x} f(u)du = \frac{1}{\pi}\int_{-\infty}^{x} \frac{1}{u^2+1}du \\ &= \frac{1}{\pi}\left[\tan^{-1}u\right]_{-\infty}^{x} = \frac{1}{\pi}\left(\tan^{-1}x + \frac{\pi}{2}\right) \\ &= \frac{1}{2} + \frac{1}{\pi}\tan^{-1}x\end{aligned}$$

5. The *rth moment of a random variable X about the mean μ*, also called the *rth central moment*, is defined as

$$m_r = E[(X-\mu)^r]$$

where $r = 0, 1, 2, \cdots$. It follows that $m_0 = 1, m_1 = 0$ and $m_2 = \sigma^2$, i.e. the second central moment or second moment about the mean is the variance.

The *rth moment of X about the origin* is defined as

$$\bar{m}_r = E[X^r]$$

where $r = 0, 1, 2, \cdots$.

Prove the relationship between these moments is given by

$$m_r = \bar{m}_r - \binom{r}{1}\bar{m}_{r-1}\mu + \cdots + (-1)^j\binom{r}{j}\bar{m}_{r-j}\mu^j + \cdots + (-1)^r\bar{m}_0\mu^r$$

Proof:

$$\begin{aligned}
m_r &= E[(X-\mu)^r] \\
&= \left[X^r - \binom{r}{1}X^{r-1}\mu + \cdots + (-1)^j\binom{r}{j}X^{r-j}\mu_j + \cdots + (-1)^{r-1}\binom{r}{r-1}X\mu^{r-1} + (-1)^r\mu^r\right] \\
&= E[X^r] - \binom{r}{1}E[X^{r-1}]\mu + \cdots + (-1)^j\binom{r}{j}E[X^{r-j}]\mu^j \\
&+ \cdots + (-1)^{r-1}\binom{r}{r-1}E[X]\mu^{r-1} + (-1)^r\mu^r \\
&= \bar{m}_r - \binom{r}{1}\bar{m}_{r-1}\mu + \cdots + (-1)^j\binom{r}{j}\bar{m}_{r-j}\mu^j \\
&+ \cdots + (-1)^{r-1}r\mu^r + (-1)^r\mu^r
\end{aligned}$$

6. Suppose that X has a Poisson distribution with mean λt; and that Y has a gamma distribution with shape parameter $= k$ and rate parameter $= \lambda$ where k is a positive integer. Show that $P[X \geq k] = P[Y \leq t]$ by showing that both the left side and the right side of this equation can be regarded as the probability of the same event in a Poisson process in which the expected number of occurrences per unit of time is λ.

   Proof: Consider the event that there are at least k occurrences between time 0 and time t. The number X of occurrences in this interval has the specified Poisson distribution, so the left side represents the probability of this event. But the event also means that the total waiting time Y until the kth event occurs is $\leq t$. We know that Y has the specified gamma distribution. Hence, the right side also expresses the probability of this same event.

7. Let N be a geometric random variable with $S_N = \{0, 1, \cdots\}$

   (a) Find $P[N > k]$.

   (b) Find the cdf of N.

   (c) Find $P[N \text{ is an even number}]$

   (d) Find $P[N = k | N \leq m]$.

Solution: Since $S_N = \{0, 1, \cdots\}$, N is the number of failures before the first success and thus the probability function is

$$P[N = k] = q^k p$$

(a)

$$P[N > k] = \sum_{j=k+1}^{\infty} q^j p = \frac{pq^{k+1}}{1-q} = q^{k+1}$$

(b)

$$F(k) = P[N \le k] = \sum_{j=0}^{k} q^j p = 1 - \sum_{j=k+1}^{\infty} q^j p = 1 - q^{k+1}$$

(c)

$$P[N \text{ is an even number}] = \sum_{\text{even } j} q^j p$$

$$= p\left(1 + q^2 + q^4 + q^6 + \cdots\right) = \frac{p}{1-q^2}$$

(d)

$$P[N = k | N \le m] = \frac{P[\{N = k\} \cap \{N \le m\}]}{P[N \le m]}$$

$$= \begin{cases} \dfrac{P[N = k]}{P[N \le m]}, & k \le m \\ 0, & k > m \end{cases}$$

$$= \begin{cases} \dfrac{q^k p}{1 - q^{m+1}}, & k \le m \\ 0, & k > m \end{cases}$$

8. A communication channel accepts an arbitrary voltage input v and outputs a voltage $Y = v + N$, where N is a Gaussian random variable with mean 0 and variance $\sigma^2 = 1$. Suppose that the channel is used to transmit binary information as follows:

   to transmit 0, input -1

   to transmit 1, input +1

   The receiver decides a 0 was sent if the voltage is negative and a 1 otherwise. Find the probability of the receiver making an error if a 0 was sent, if a 1 was sent.

   Solution: The voltage Y has a Gaussian distribution with mean $\mu = v$ and standard deviation 1. Thus its density is

$$f_Y(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-v)^2}$$

where $v = -1$ when a 0 was sent and $v = +1$ when a +1 was sent.

When $v = -1$, the receiver will send a +1 if $x \geq 0$ and thus an error would be made. This probability is

$$P[x \geq 0] = \frac{1}{\sqrt{2\pi}} \int_0^\infty e^{-\frac{1}{2}(x+1)^2} dx \quad (1)$$

When $v = +1$, the receiver will send a 0 if $x \leq 0$ and thus an error would be made. This probability is

$$P[x \leq 0] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^0 e^{-\frac{1}{2}(x-1)^2} dx \quad (2)$$

Eq: (1) and (2) are the probabilities of the tail of the pdf. There are no closed-form expression. Numerical calculations can be used (see the sample R code) and we obtain:

$$P[x \geq 0 | v = -1] = 0.1586, \quad P[x \leq 0 | v = +1] = 0.1586$$

9. Suppose that we observe the times S_m that elapses until the occurrence of the m-th event. The times $X_1, X_2, \cdots, X_m$ between events are exponential random variables, so we must have

$$S_m = X_1 + X_2 + \cdots + X_m$$

S_m is the m-Erland random variable.

(a) Find the cdf of S_m by integration of the pdf. Hint: Use integration by parts.

(b) Show that the derivative of the cdf obtained in (a) gives the pdf of an m-Erland random variable.

Solution: (a) The pdf of S_m is

$$f(x) = \begin{cases} \frac{\lambda^m}{\Gamma(m)} x^{m-1} e^{-\lambda x}, & x > 0 \\ 0, & \text{otherwise} \end{cases}$$

Thus, the cdf of S_m can be obtained by integrating $f(x)$ as follows:

$$F(t) = P[S_m \leq t] = \int_0^t \frac{\lambda^m}{\Gamma(m)} x^{m-1} e^{-\lambda x} dx$$
$$= \frac{\lambda^m}{\Gamma(m)} \int_0^t x^{m-1} e^{-\lambda x} dx$$
$$= \frac{\lambda^m}{\Gamma(m)} \left[\left(-\frac{1}{\lambda} \right) e^{-\lambda x} x^{m-1} \bigg|_0^t + \frac{1}{\lambda} \int_0^t e^{-\lambda x} (m-1) x^{m-2} dx \right]$$
$$= \frac{\lambda^m}{\Gamma(m)} \left[-\frac{1}{\lambda} e^{-\lambda t} t^{m-1} + \frac{m-1}{\lambda} \int_0^t e^{-\lambda x} x^{m-2} dx \right]$$
$$= -\frac{(\lambda t)^{m-1}}{\Gamma(m)} e^{-\lambda t} + \frac{\lambda^{m-1}(m-1)}{\Gamma(m)} \left[\left(-\frac{1}{\lambda} e^{-\lambda x} \right) x^{m-2} \bigg|_0^t + \frac{m-2}{\lambda} \int_0^t e^{-\lambda x} x^{m-3} dx \right]$$
$$= -\frac{(\lambda t)^{m-1}}{(m-1)!} e^{-\lambda t} + \frac{\lambda^{m-1}}{(m-2)!} \left[-\frac{1}{\lambda} e^{-\lambda t} t^{m-2} + \frac{m-2}{\lambda} \int_0^t e^{-\lambda x} x^{m-3} dx \right]$$
$$= -\frac{(\lambda t)^{m-1}}{(m-1)!} e^{-\lambda t} - \frac{(\lambda t)^{m-2}}{(m-2)!} e^{-\lambda t} + \frac{\lambda^{m-2}}{(m-3)!} \int_0^t e^{-\lambda x} x^{m-3} dx$$
$$= -\frac{(\lambda t)^{m-1}}{(m-1)!} e^{-\lambda t} - \frac{(\lambda t)^{m-2}}{(m-2)!} e^{-\lambda t} - \cdots + \frac{\lambda^2}{1!} \int_0^t e^{-\lambda x} x dx$$
$$= -\frac{(\lambda t)^{m-1}}{(m-1)!} e^{-\lambda t} - \frac{(\lambda t)^{m-2}}{(m-2)!} e^{-\lambda t} - \cdots + \lambda^2 \left[-\frac{1}{\lambda} e^{-\lambda x} x \bigg|_0^t + \frac{1}{\lambda} \int_0^t e^{-\lambda x} dx \right]$$
$$= -\frac{(\lambda t)^{m-1}}{(m-1)!} e^{-\lambda t} - \frac{(\lambda t)^{m-2}}{(m-2)!} e^{-\lambda t} - \cdots - (\lambda t) e^{-\lambda t} + \lambda \int_0^t e^{-\lambda x} dx$$
$$= -\frac{(\lambda t)^{m-1}}{(m-1)!} e^{-\lambda t} - \frac{(\lambda t)^{m-2}}{(m-2)!} e^{-\lambda t} - \cdots - (\lambda t) e^{-\lambda t} + \lambda \left(-\frac{1}{\lambda} e^{-\lambda x} \right) \bigg|_0^t$$
$$= -\frac{(\lambda t)^{m-1}}{(m-1)!} e^{-\lambda t} - \frac{(\lambda t)^{m-2}}{(m-2)!} e^{-\lambda t} - \cdots - (\lambda t) e^{-\lambda t} - e^{-\lambda t} + 1$$
$$= 1 - \sum_{k=0}^{m-1} \frac{(\lambda t)^k}{k!} e^{-\lambda t}$$

The cdf of S_m can also be found by using its close interrelation with the exponential and Poisson random variables. Let $N(t)$ be the Poisson random variable for the number of event in t seconds. Note that the mth event occurs before time t, that is, $S_m \leq t$, if and only if m or more events occur in t seconds, namely $N(t) \geq m$. The reasoning goes as follows. If the mth event has occurred before time t, then it follows that m or more events will occur in time t. On the other hand, if m or more events occur in time t, then it follows that the mth event occurred by time t. Thus

$$F(t) = P[S_m \leq t] = P[N(t) \geq m] = 1 - \sum_{k=0}^{m-1} \frac{(\lambda t)^k}{k!} e^{-\lambda t}$$

In general, the cdf of the gamma random variable does not have a closed-form expression. But the special case, the m-Erlang random variable, does have have a closed-form expression.

(b) Differentiate $F(t)$, we obtain

$$\frac{d}{dt}F(t) = \frac{d}{dt}\left[1 - \left(1 + \lambda t + \frac{(\lambda t)^2}{2!} + \frac{(\lambda t)^3}{3!} + \cdots + \frac{(\lambda t)^{m-1}}{(m-1)!}\right) e^{-\lambda t}\right]$$

$$= 0 - \left(0 + \lambda + \frac{\lambda^2 2t}{2!} + \frac{\lambda^3 3t^2}{3!} + \cdots + \frac{\lambda^{m-1}(m-1)t^{m-2}}{(m-1)!}\right) e^{-\lambda t}$$

$$- \left(1 + \lambda t + \frac{(\lambda t)^2}{2!} + \frac{(\lambda t)^3}{3!} + \cdots + \frac{(\lambda t)^{m-1}}{(m-1)!}\right) (-\lambda e^{-\lambda t})$$

$$= - \left(\lambda + \frac{\lambda^2 t}{1!} + \frac{\lambda^3 t^2}{2!} + \cdots + \frac{\lambda^{m-1} t^{m-2}}{(m-2)!}\right) e^{-\lambda t}$$

$$- \left(-\lambda - \frac{\lambda^2 t}{1!} - \frac{\lambda^3 t^2}{2!} - \cdots - \frac{\lambda^m t^{m-1}}{(m-1)!}\right) e^{-\lambda t}$$

$$= \frac{\lambda^m}{\Gamma(m)} t^{m-1} e^{-\lambda t}$$