Mathematics 50
Probability and Statistical Inference
Winter 1999

1. Univariate Continuous Random Variables

(continued)

Week 2
January 11-15

1. Moments

Random variable is completely defined by its distribution or density function. Sometimes it is convenient to characterize RV by a number or at least a few numbers rather than a function. Moments fulfill this role. The most important moments are mean and variance.

1.1. Mean and variance

Definition 1.1. *The mean of CRV is defined as*

$$mean = \int_{-\infty}^{\infty} x f(x) dx.$$

Sometimes mean is called expected value, and the notation is used, $mean = E(X)$. By convention, we use μ ('mu') to denote the mean.

Mean may not exist – show that it does not exist for Cauchy distribution.

Recall, for a discrete RV the mean is defined via the sum,

$$\sum p_i x_i$$

where x_i is a value X may take and $p_i = \Pr(X = x_i)$ the corresponding probability. The sum replaces the integral.

Mean is the parameter of location.

Definition 1.2. *Variance is defined as*

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

Standard Deviation (SD) is

$$SD = \sqrt{variance}.$$

The advantage of SD over variance is that it has the same units of measurement as the random variable (what is the unit of measurement of the variance?) Often, variance is denoted as σ^2 ('sigma squared'). Then SD=σ.

Alternatively we can write

$$\begin{aligned}\sigma^2 &= E(X^2) - E^2(X) \\ &= E(X^2) - \mu^2.\end{aligned}$$

Mean, mode and median are parameters of location, i.e. they determine 'the center' of the distribution. Variance determines how random variable spreads around the mean. Show that if variance=0 then RV degenerates (becomes constant).

Alternative formula for variance:

$$\sigma^2 = \int_{-\infty}^{\infty} x^2 f(x)dx - \mu^2. \tag{1.1}$$

Proof. Since

$$(x-\mu)^2 = x^2 - 2x \times \mu + \mu^2,$$

we obtain

$$\begin{aligned}\sigma^2 &= \int_{-\infty}^{\infty} (x-\mu)^2 f(x)dx \\ &= \int_{-\infty}^{\infty} x^2 f(x)dx - 2\mu \times \int_{-\infty}^{\infty} x f(x)dx + \mu^2 \int_{-\infty}^{\infty} f(x)dx \\ &= \int_{-\infty}^{\infty} x^2 f(x)dx - 2\mu \times \mu + \mu^2 \\ &= \int_{-\infty}^{\infty} x^2 f(x)dx - \mu^2.\end{aligned}$$

1.1.1. Uniform distribution

Mean:

$$\begin{aligned}E(X) &= \int_{-\infty}^{\infty} x f(x)dx = \int_a^b x f(x)dx \\ &= \frac{1}{b-a}\int_a^b x dx = \frac{1}{b-a}\left(\frac{1}{2}x^2\Big|_a^b\right) \\ &= \frac{1}{b-a}\left(\frac{1}{2}b^2 - \frac{1}{2}a^2\right) = \frac{1}{2}\frac{1}{b-a}(b-a)(a+b) \\ &= \frac{1}{2}(a+b)\end{aligned}$$

i.e.

$$\text{mean of uniform RV} = \frac{1}{2}(a+b).$$

Variance: Using alternative formula for variance we obtain

$$\begin{aligned}E(X^2) &= \int_{-\infty}^{\infty} x^2 f(x)dx = \frac{1}{b-a}\int_a^b x^2 dx \\ &= \frac{1}{b-a}\left(\frac{1}{3}b^3 - \frac{1}{3}a^3\right) = \frac{1}{3}\frac{1}{b-a}(b-a)(a^2+ab+b^2) \\ &= \frac{1}{3}(a^2+ab+b^2)\end{aligned}$$

Now the variance is

$$
\begin{aligned}
&\int_{-\infty}^{\infty} x^2 f(x)dx - (E(X))^2 = \frac{1}{3}(a^2+ab+b^2) - \frac{1}{4}(a+b)^2 \\
=\ & \frac{1}{12}(4a^2+4ab+4b^2-3a^2-6ab-3b^2) = \frac{1}{12}(a^2-2ab+b^2) \\
=\ & \frac{1}{12}(b-a)^2
\end{aligned}
$$

$$\text{variance of uniform RV} = \frac{1}{12}(b-a)^2$$

$$\text{SD} = \frac{1}{2\sqrt{3}}(b-a).$$

1.2. Central and noncentral moments

We can find the expected value of any power of CRV,

$$\mu_k = \int_{-\infty}^{\infty} x^k f(x)dx$$

it is called the kth (*non-central*) moment, if exists ($k = 1, 2, 3 \ldots$).

$$\nu_k = \int_{-\infty}^{\infty} (x-\mu)^k f(x)dx$$

is called the kth *central* moment, if exists, where $\mu = E(X) = \mu_1$, the mean (the first moment). $\nu_2 = \sigma^2$, the variance (the second central moment).

1.3. Moment generating function

Moments calculation is a tedious task. It was invented Moment Generating Function (MGF)

$$M(t) = \int_{-\infty}^{\infty} e^{tx} f(x)dx,$$

to compute moments easier. MGF may not exist for any t (integral may diverge), but we actually need $M(t)$ in the neighborhood of zero, see below.

Theorem 1.3. *The kth moment is the kth derivative of MGF at zero:*

$$\mu_k = \left.\frac{d^{(k)}M(t)}{d^k t}\right|_{t=0}$$

Proof (sketch). We have

$$\frac{d^{(k)}M(t)}{d^k t} = \int_{-\infty}^{\infty} x^k e^{tx} f(x)dx$$

so that

$$\left.\frac{d^{(k)}M(t)}{d^k t}\right|_{t=0} = \int_{-\infty}^{\infty} x^k f(x)dx = \mu_k.$$

One can show that distribution is uniquely defined by all its moments.

2. Exponential distribution

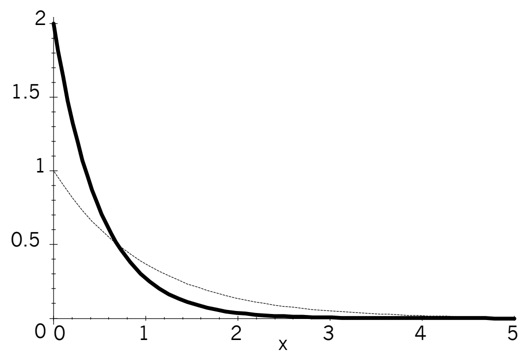
is defined by the density,

$$f(x) = \lambda e^{-\lambda x}, \qquad x > 0$$

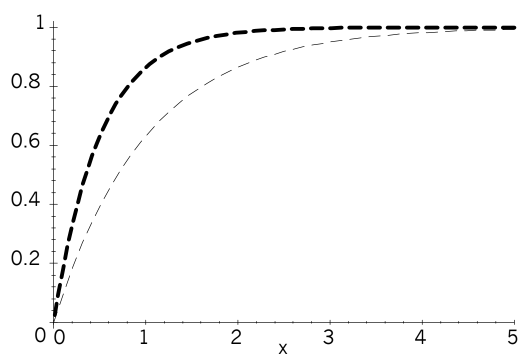
where *parameter* λ (positive) is called 'rate failure'. The support of the exponential distribution is all positive numbers. Often distributions have parameters, i.e. in fact we introduce a family of distributions, and each member of the family is specified by parameter.

The distribution function is

$$F(x) = \int_{-\infty}^{x} \lambda e^{-\lambda t} dt = \int_{0}^{x} \lambda e^{-\lambda t} dt$$
$$= -e^{-\lambda t}\Big|_0^x = 1 - e^{-\lambda x}.$$



Two densities of exponential distributions with different rates, $\lambda = 1$ (dashed), $\lambda = 2$ (solid). Rate controls the speepness of the curve.



Distribution functions for $\lambda = 1$ (dashed) and $\lambda = 2$ (solid).

Exponential distribution is used to model *waiting* or *lifetime time* (that is why sometimes t is used instead of x). A characteristic feature of the exponential distribution is that its density is a decreasing function, i.e. the maximum density is at zero and then it gradually drops. We shall learn later that λ characterizes how frequently events happen within one unit of time. Then, $1 - e^{-\lambda x}$ gives the probability that at least one event happens in the time frame $(0, x)$.

Problem. Find MGF of exponential distribution.

Solution. We have

$$M(t) = \int_0^{\infty} \lambda e^{-\lambda x} e^{xt} dx = \lambda \int_0^{\infty} e^{-(\lambda - t)x} dx.$$

We observe that

$$\int_0^\infty e^{-(\lambda-t)x}dx = \frac{1}{\lambda-t}$$

because $(\lambda - t)$ can be treated as the rate failure. Also, the integral diverge if $\lambda - t < 0$, so t must be small enough. Hence, the MGF for exponential distribution is

$$M(t) = \frac{\lambda}{\lambda-t} = \lambda(\lambda-t)^{-1}.$$

According to the above Theorem, the first and the second moments are calculated as follows:

$$\begin{aligned}
E(X) &= \left.\frac{dM(t)}{dt}\right|_{t=0} = \lambda\left.(\lambda-t)^{-2}\right|_{t=0} \\
&= \frac{1}{\lambda}, \\
E(X^2) &= \left.\frac{d^2M(t)}{dt^2}\right|_{t=0} = 2\lambda\left.(\lambda-t)^{-3}\right|_{t=0} = \frac{2}{\lambda^2}
\end{aligned}$$

Using the alternative formula for the variance we obtain

$$var(X) = E(X^2) - E^2(X) = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}.$$

For any moment

$$E(X^k) = k!\lambda^{-k}.$$

Example 1. Let us assume an asteroid hits the Earth, following exponential distribution, every 1000 years (the mean). What is the probability that it hits the Earth during you life-span (say 85 years).

Solution. Since $E(X) = 1000$ we obtain $\lambda = 1/1000$. Therefore, the probability that asteroid hits the Earth when you are alive is

$$1 - e^{-85/1000} = 1 - 0.91851 = 0.0815.$$

Example 2. Bill said that he will call after 10 am. Assume that Pr(Bill calls from 10 to $10+x$) follows exponential distribution $1 - e^{-\lambda x}$ with parameter $\lambda = .1$ where x is measured in minutes. What decision gives a better probability to talk with Bill: (a) wait first 10 minutes, (b) wait the call from 10:10 to 11:00?

Solution. The distribution function is

$$F(x) = 1 - e^{-.1x}.$$

The first probability (a) is

$$\begin{aligned}
\Pr(X &< 10) = F(10) \\
&= 1 - \exp(-10 \times 0.1) \\
&= 1 - 1/e \\
&= .63212.
\end{aligned}$$

The second probability (b) is

$$\begin{aligned}\Pr(10 &< X < 60) \\ &= F(60) - F(10) \\ &= (1 - e^{-.1\times 60}) - (1 - e^{-.1\times 10}) \\ &= e^{-.1\times 10} - e^{-.1\times 60} \\ &= e^{-1} - e^{-6} \\ &= .3654.\end{aligned}$$

Therefore, it is much better to wait the call first 10 minutes than to wait from 10:10 to 11:00. Surprised? How can you explain this? What assumption drives the answer?

Exponential distribution has no memory.

Continuing the above example, what is the probability that if Bill did not call in the interval $10 + x$ he does not call in the interval $10 + x + t$. This probability is a conditional probability

$$\begin{aligned}\Pr(X &> 10 + x + t | X > 10 + x) \\ &= \frac{\Pr(X > 10 + x + t \text{ AND } X > x)}{\Pr(X > 10 + x)} \\ &= \frac{\Pr(X > 10 + x + t)}{\Pr(X > 10 + x)} \\ &= \frac{1 - \Pr(X \le 10 + x + t)}{1 - \Pr(X \le 10 + x)} \\ &= \frac{1 - (1 - e^{-\lambda(x+t)})}{1 - (1 - e^{-\lambda x})} \\ &= \frac{e^{-\lambda(x+t)}}{e^{-\lambda x}} \\ &= e^{-\lambda t}.\end{aligned}$$

Conclusion. The probability does not depend on the history (memoryless), i.e. at what time (x) it is calculated. It may be OK for this example, but it is not OK, e.g. to model human lifetime. Otherwise, the probability to live 10 more years is the same for 16 years old boy as for 85 years old man.

Is the exponential distribution valid to model the time to failure of an electronic component?

3. Gamma distribution

is a generalization of exponential distribution. However, unlike exponential distribution it has memory. Gamma distribution is tightly related to Poisson distribution.

The density is defined as

$$f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, \qquad x > 0$$

where Γ is the *Gamma function*. It is defined as the integral,

$$\Gamma(\alpha) = \int_0^\infty u^{\alpha-1} e^{-u} du.$$

If α is integer then

$$\Gamma(\alpha) = 1 \cdot 2 \cdots (\alpha - 1) = (\alpha - 1)!$$

Also, it is worthwhile to remember that

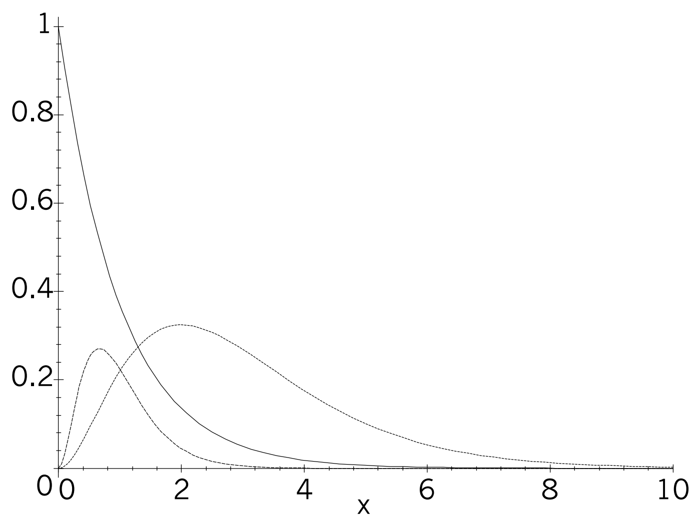
$$\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1),$$

$$\Gamma\left(\frac{1}{2}\right) = \frac{\pi}{2}.$$

Also,

$$\Gamma(1) = \Gamma(2) = 1.$$

Notice Gamma distribution has two parameters, λ and α (both are positive, two-parameter family). Sometimes, to show the dependence on parameters, we write $f(x; \alpha, \lambda)$. When $\alpha = 1$ we come to the exponential distribution, i.e. Gamma distribution is a richer distribution. Gamma distribution also is used to model the waiting time.



Three Gamma-densities: $\alpha = 3, \lambda = 1$ (dashed) $\alpha = 1, \lambda = 1$ (solid, exponential distribution); $\alpha = 3, \lambda = 3$ (dotted).

The moment generating function for Gamma distribution

$$\begin{aligned} M(t) &= \frac{\lambda^\alpha}{\Gamma(\alpha)} \int x^{\alpha-1} e^{-\lambda x} e^{tx} dx = \frac{\lambda^\alpha}{\Gamma(\alpha)} \int x^{\alpha-1} e^{-(\lambda - t)x} dx \\ &= \frac{\lambda^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha)}{(\lambda - t)^\alpha} = \frac{1}{(1 - t/\lambda)^\alpha}. \end{aligned}$$

The mean and the variance of the Gamma distribution

$$E(X) = \frac{\alpha}{\lambda}, \qquad var(X) = \frac{\alpha}{\lambda^2}.$$

Relationship to Poisson distribution. If events happen with the mean rate λ per unit time according to the Poisson distribution, then the distribution of the waiting time until the αth event has Gamma-distribution with parameter λ and α.

Example. You work at the college telephone board as an operator and you are late by 5 minutes. The rule is that you will be fired if you miss 3 calls in a row. Assuming calls arrive at $\lambda = 2$ calls per minute according to the Poisson distribution, what is the probability you will be fired.

Solution. Since $\lambda = 2$ and $\alpha = 3$ the probability is

$$\Pr(X \leq 5) = \frac{\lambda^\alpha}{\Gamma(\alpha)} \int_0^5 x^{\alpha-1} e^{-\lambda x} dx$$

or

$$\begin{aligned} \Pr(X &\leq 5) \\ &= \frac{2^3}{3!} \int_0^5 x^2 e^{-2x} dx \\ &= \frac{2^3}{2!} \int_0^5 x^2 e^{-2x} dx \\ &= .99723 \\ &\simeq 1, \end{aligned}$$

almost one. Use integration by parts twice,

$$\int u dv = uv - \int v du$$

to calculate

$$\int_0^5 x^2 e^{-2x} dx.$$

4. Properties of the mean

Recall

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx.$$

Mean (expected value) of a function, $g(X)$:

$$E(g(X)) = \int_{-\infty}^{\infty} g(x) f(x) dx.$$

Properties of the mean:

1. $E(c) = c$
2. $E(cg(X)) = cE(g(X))$
3. $E(g_1(X) \pm g_2(X)) = E(g_1(X)) \pm E(g_2(X))$
4. Generally,

$$\begin{aligned} & E\left(c_1 g_1(X) + c_2 g_2(X) + ... + c_n g_n(X)\right) \\ = \; & c_1 E(g_1(X)) + c_2 E(g_2(X)) + ... + c_n E(g_n(X)). \end{aligned}$$

Expected value of linear combination is linear combination of expected values.

These facts follow from the following integral property

$$
\begin{aligned}
& \int (c_1 g_1(x) + c_2 g_2(x)) f(x) dx \\
= \; & c_1 \int g_1(x) f(x) dx + c_2 \int g_2(x) f(x) dx.
\end{aligned}
$$

Problem. Find $E(X - 2e^X)$ where $X \sim U(a, b)$.

Solution. By rule #3 we have $E(X - 2e^X) = E(X) - 2E(e^X)$. But we know that

$$E(X) = \frac{1}{2}(a+b).$$

As follows from the definition,

$$
\begin{aligned}
E(e^X) &= \frac{1}{b-a} \int_a^b e^x dx \\
&= \frac{e^b - e^a}{b-a}.
\end{aligned}
$$

Finally,

$$E(X - e^X) = \frac{1}{2}(a+b) - 2\frac{e^b - e^a}{b-a}.$$

5. Function of random variable

Let X be a random variable and g any function. What is the distribution of $g(X)$, i.e. how can the distribution of $g(X)$ be expressed via the distribution of X?

We start with the simplest case, linear function (transformation)

$$g(x) = ax + b$$

with positive slope ($a > 0$). Let

$$Y = g(X) = aX + b.$$

Denote F_X – the distribution and f_X – the density of X. Then the distribution function of Y (using the definition of df) is

$$
\begin{aligned}
F_Y(y) &= \Pr(Y \le y) \\
&= \Pr(aX + b \le y) \\
&= \Pr(aX \le y - b) \\
&= \Pr\left(X \le \frac{y-b}{a}\right) \\
&= F_X\left(\frac{y-b}{a}\right)
\end{aligned}
$$

Thus, the distribution of $Y = aX + b$ is

$$F_Y(y) = F_X\left(\frac{y-b}{a}\right). \tag{5.1}$$

Take the derivative and obtain the density:

$$f_Y(y) = \frac{1}{a} f_X\left(\frac{y-b}{a}\right).$$

In general case, i.e. when a may be negative but not zero, we can prove that the density of Y is

$$f_Y(y) = \frac{1}{|a|} f_X\left(\frac{y-b}{a}\right).$$

Example. X has exponential distribution with rate λ. Find the distribution and the density functions of $Y = aX$ with $a > 0$. We have

$$F_X(x) = 1 - e^{-\lambda x}.$$

Using (5.1) we obtain that

$$F_Y(y) = 1 - e^{-\lambda(y/a)}$$

also an exponential distribution with the rate λ/a. The density is

$$f_Y(y) = \frac{\lambda}{a} e^{-(\lambda/a)y}.$$

We learned: linear transformation of an exponential distribution is again an exponential distribution.

General case – any transformation:

$$Y = g(X).$$

What is the distribution and density of Y? Denote g^{-1} the inverse function ($g(g^{-1}(y)) = y$). Example: if $y = \ln(x)$ then $g(x) = \ln(x)$ and $g^{-1}(y) = e^y$. We assume g is an increasing function , i.e. $dg/dx > 0$.

As in linear case

$$\begin{aligned}
F_Y(y) &= \Pr(Y \leq y) \\
&= \Pr(g(X) \leq y) \\
&= \Pr(X \leq g^{-1}(y)) \\
&= F_X\left(g^{-1}(y)\right).
\end{aligned}$$

Therefore, the distribution function of $g(X)$ is

$$F_X\left(g^{-1}(y)\right) = \Pr(Y \leq y), \tag{5.2}$$

and the density function (take derivative)

$$f_Y(y) = \frac{dg^{-1}(y)}{dy} f_X\left(g^{-1}(y)\right).$$

In general case, when g is monotone (but still $g' \neq 0$) the formula is

$$f_Y(y) = \left|\frac{dg^{-1}(y)}{dy}\right| f_X\left(g^{-1}(y)\right). \tag{5.3}$$

Example (continued). Find the distribution and the density of the log exponential distribution.

Solution. We have $y = \ln(x) = g(x)$ and $g^{-1}(y) = e^y$.Therefore, using formula (5.2) the distribution of Y is

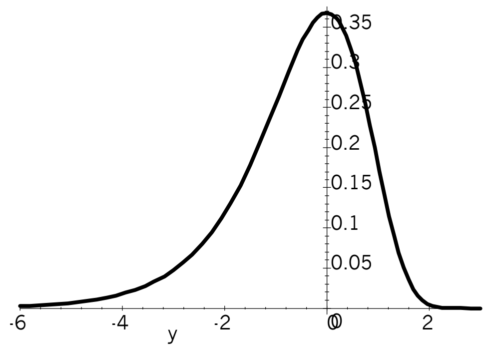
$$1 - e^{-\lambda e^y}.$$

Since

$$\frac{dg^{-1}}{dy} = \frac{de^y}{dy} = e^y,$$

using formula (5.3), the density becomes

$$\lambda e^y e^{-\lambda e^y} = \lambda e^{y-\lambda e^y}.$$



The density function of the log exponential distribution
for the unit rate failure. The support is (-∞, $+\infty$).

6. Homework (due January 20, Wednesday)

Maximum number of points is 37.

1. (4 points). Double exponential distribution (sometimes called *Laplace* distribution) is defined by the density $.5\exp(-|x|)$. Find MGF and all moments for this distribution.

Solution. By definition, for t in the neighborhood of zero

$$\begin{aligned} M(t) &= \frac{1}{2}\int_{-\infty}^{\infty} e^{tx-|x|}dx = \frac{1}{2}\left(\int_{-\infty}^{0} e^{tx-|x|}dx + \int_{0}^{\infty} e^{tx-|x|}dx\right) = \frac{1}{2}\left(\int_{-\infty}^{0} e^{tx+x}dx + \int_{0}^{\infty} e^{tx-x}dx\right) \\ &= \frac{1}{2}\left(\int_{-\infty}^{0} e^{(t+1)x}dx + \int_{0}^{\infty} e^{-(1-t)x}dx\right) = \frac{1}{2}\left(\frac{1}{t+1}e^{(t+1)x}\Big|_{-\infty}^{0} - \frac{1}{1-t}e^{-(1-t)x}\Big|_{0}^{\infty}\right) \\ &= \frac{1}{2}\left(\frac{1}{t+1} + \frac{1}{1-t}\right). \end{aligned}$$

To find the kth moment we need the kth derivative of $M(t)$ at $t = 0$. It is easy to check that

$$\frac{d^k}{dt^k}\left(\frac{1}{t+1}\right) = \frac{d^k}{dt^k}(t+1)^{-1} = (-1)^k k!(1+t)^{-(k+1)}$$

and similarly

$$\frac{d^k}{dt^k}\left(\frac{1}{1-t}\right) = \frac{d^k}{dt^k}(1-t)^{-1} = k!(1-t)^{-(k+1)}$$

Combining, we obtain

$$\frac{1}{2}\frac{d^k}{dt^k}\left(\frac{1}{1-t^2}\right)\Bigg|_{t=0}=\begin{cases}0 & if \quad k \text{ is odd}\\ k! & if \quad k \text{ is even}\end{cases}.$$

The answer: all odd moments are zero (because *Laplace* distribution is symmetric around zero) and the kth even moment is $k!$.

2. (3 points). Prove that variance does not change for any additive shift, i.e. $var(X) = var(X + c)$ where c is a constant. Does it hold if c is not a constant but a RV?

Solution. Denote $Y = X + c$, then $\mu_Y = E(X + c) = E(X) + c = \mu_X + c$. The variance is

$$var(Y)=\int_{-\infty}^{\infty}(x+c-\mu_Y)^2 f_X(x)dx=\int_{-\infty}^{\infty}(x-\mu_X)^2 f_X(x)dx=var(X).$$

It is not true if c is not a constant.

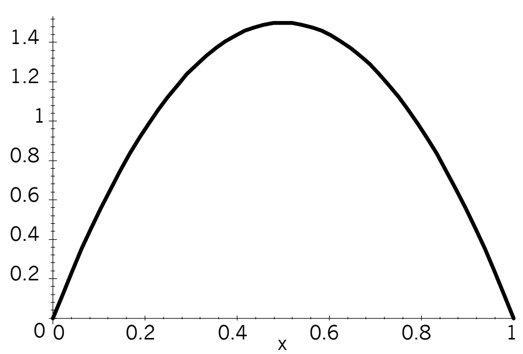
3. (2 points). Show that all odd central moments of a symmetric distribution are zero.

Solution. If μ is the mean (mode) of a symmetric distribution then $f(\mu - x) = f(\mu + x)$. The kth central moment is

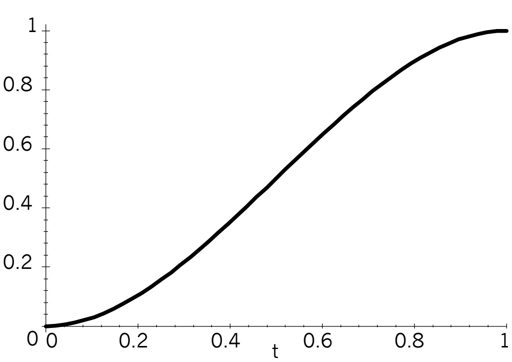
$$\begin{aligned}\mu_k &= \int_{-\infty}^{\infty}(x-\mu)^k f(x)dx=\int_{-\infty}^{\mu}(x-\mu)^k f(x)dx+\int_{\mu}^{\infty}(x-\mu)^k f(x)dx\\ &= \int_{-\infty}^{0} z^k f(\mu+z)dz+\int_0^{\infty} z^k f(\mu+z)dz=-\int_0^{\infty} z^k f(\mu-z)dz+\int_0^{\infty} z^k f(\mu+z)dz\\ &= -\int_0^{\infty} z^k f(\mu-z)dz+\int_0^{\infty} z^k f(\mu-z)dz\\ &= 0.\end{aligned}$$

4. (5 points). CRV X takes values on the interval $(0, 1)$ with the density defined as $f_X(x) = cx(1 - x)$. Find constant c. Find distribution function $F_X(x)$ and sketch its graph along with the density. Find the mean and the variance.

Solution. Constant c is found from $\int_0^1 f_X(x)dx = 1$. But $\int_0^1 x(1-x)dx = \int_0^1 xdx - \int_0^1 x^2dx = 0.16667$ so that $c = 1/0.16667 = 6$. The d.f. is $t^2(3 - 2t)$



Density $6x(1-x)$, it is 0 outside of $(0, 1)$.



d.f. $x^2(3 - 2x)$

Mean:

$$\mu = 6\int_0^1 x^2(1-x)dx = .5.$$

But

$$6\int_0^1 x^3(1-x)dx = .3,$$

so that the variance

$$\sigma^2 = .3 - (.5)^2 = .05.$$

5. (2 points). Solve problem 2.5.45 from Rice (p. 65).

Solution. The lifetime has the distribution $F(x) = 1 - e^{-.1x}$.

a. $\Pr(X < 10) = 1 - e^{-1} = 0.63212$

b. $\Pr(5 < X < 15) = \Pr(X < 15) - \Pr(X < 5) = e^{-.5} - e^{-1.5} = 0.3834$

c. t is the solution to $\Pr(X > t) = 0.01$, i.e. $e^{-.1t} = 0.01$. Therefore, $t = -\ln(0.01)/0.1 = 46.052$.

6. (4 points). The robber knows that the alarm starts as soon he gets into the bank. The robber also knows that he needs 2 minutes to finish off his business. Assuming that the time of police arrival follows exponential distribution with expected value 0.25 what is the probability that the robber will not be caught?

Solution. Let X denote the time of police arrival after the robber gets in. X follows exponential distribution with the rate λ. Since the expected value of X, $E(X) = .25$, and since we know that for the exponential distribution $E(X) = 1/\lambda$ we find $\lambda = 1/.25 = 4$. The distribution function of X is $F(x) = \Pr(X \leq x) = 1 - e^{-\lambda x} = 1 - e^{-4x}$. The probability not to be caught is $\Pr(X > 2) = 1 - F(2) = e^{-4\times 2} = e^{-8} = 0.0033$.

7. (2 points). Define failure rate of a random variable as $f(x)/(1-F(x))$. Show that the failure rate is constant for exponential distribution.

Solution. We have $f(x) = \lambda \exp(-\lambda x)$ and $F(x) = 1 - \exp(-\lambda x)$. Then

$$\frac{f(x)}{1-F(x)} = \frac{\lambda e^{-\lambda x}}{e^{-\lambda x}} = \lambda.$$

8. (5 points). Bill said that he will call after 10 am. Assume that Pr(Bill calls from 10 to 10+x) follows a Gamma distribution with parameters $\lambda = .1$ and $\alpha = 2$ where x is measured in minutes. What decision gives better probability to talk with Bill: (a) wait first 20 minutes, (b) wait the call from 10:20 until you die (forever)? Show the work.

Solution. Let X denote time collapsed after 10 when Bill calls. The density of X is $.005x\exp(-.1x)$. The first probability is, after integration by parts, is

$$\Pr(X < 20) = 0.01 \int_0^{20} x e^{-0.1x} dx = .594$$

the second probability is

$$\Pr(X > 20) = 1 - \Pr(X < 20) = 1 - .594 = 0.406.$$

The first decision is better.

9. (4 points). The mean and the variance of a Gamma distribution are equal 2. Find the mode. Find its MGF and all moments.

Solution. The mean of a Gamma distribution is α/λ and the variance is α/λ^2. We have $\alpha/\lambda = \alpha/\lambda^2 = 2$ that leads to $\lambda = 1$ and $\alpha = 2$.

The mode is where the density takes the maximum. In our case the density is $f(x) = xe^{-x}$ and its derivative is

$$\frac{df}{dx} = e^{-x}(1 - x^2).$$

It takes zero when $x = 1$, i.e. mode=1. The MGF is $M(t) = (1-t)^{-2}$. The kth derivative is $(k+1)!(1-t)^{-(k+2)}$ so that the kth moment is $(k+1)!$.

10. (3 points). Define a function as $g(c) = E(X-c)^2$. Prove that the minimum of this is the variance.

Solution. We have

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

Its minimum attains where $g' = 0$. But

$$\frac{dg}{dc} = 2\int_{-\infty}^{\infty} (x-c)f(x)dx = 0$$

implies

$$\int_{-\infty}^{\infty} (x-c)f(x)dx = \int_{-\infty}^{\infty} xf(x)dx - c\int_{-\infty}^{\infty} f(x)dx = E(X) - c = 0,$$

which gives $E(X) = c$.

11. (3 points). Let $X \sim U(0,1)$. Find the density of $Y = 2\ln(X+1)$ and sketch its graph. Find the mean and the variance of Y.

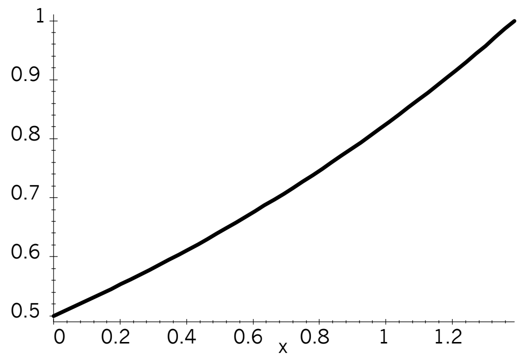
Solution. Function g is defined by $x = 2\ln(u+1)$ and the inverse function is $g^{-1}(x) = e^{x/2} - 1$. We have

$$\frac{dg^{-1}}{dx} = \frac{1}{2}e^{x/2}$$

so that the density of X is

$$f_X(x) = \begin{cases} \frac{1}{2}e^{x/2} & if \quad 0 < x < 2\ln 2 \\ 0 & otherwise \end{cases},$$

because the density of U is 1.



Density function of $2\ln(X+1)$. Outside of interval (0,1.386) the density is zero.

The mean is

$$E(X) = 2\int_0^1 \ln(x+1)dx.$$

Change the variable $y = \ln(x+1)$, then integration by part gives

$$E(X) = 2\int_0^1 \ln(x+1)dx = 2\int_0^{\ln 2} ye^y dy = 0.77259.$$

We use the formula $var(X) = E(X^2) - E^2(X)$. For this we need the integral $\int_0^1 \ln^2(x+1)dx$. Then

$$E(X^2) = 4\int_0^{\ln 2} \ln^2(x+1)dx = 4\int_0^{\ln 2} y^2 e^y dy.$$

Integration by parts gives

$$\int_0^{\ln 2} y^2 e^y dy = y^2 e^y \Big|_0^{\ln 2} - 2\int_0^{\ln 2} y e^y dy = 2\ln^2 2 - 2\int_0^{\ln 2} y e^y dy.$$

But

$$\int_0^{\ln 2} y e^y dy = 2\ln 2 - \int_0^{\ln 2} e^y dy = 2\ln 2 - 1$$

so that

$$E(X^2) = 4\left(2\ln^2 2 - 2(2\ln 2 - 1)\right) = .75327$$

Finally, find the variance

$$var(X) = .75327 - 0.77259^2 = 0.15637.$$