Continuous Probability Distributions, chapter 5

Grethe Hystad

October 29, 2013

The exponential distribution, chapter 5.4

- The Probability over intervals of constant length decreases as the intervals moves furher to the right.
- The life lengths of batteries.
- The weekly rainfall total for a section of a state.
- Time between successive hurricanes.

The exponential distribution function

Theorem

The exponential distribution:

$$f(x) = \begin{cases} \frac{1}{\theta} e^{-\frac{x}{\theta}}, & x \geq 0; \\ 0, & \textit{elsewhere} \end{cases}$$

for $\theta > 0$.

$$E(X) = \theta \;\; \textit{and} \;\; V(X) = \theta^2.$$

The parameter $\theta > 0$ is a constant that determines the rate at which the curve decreases. The cumulative distribution function:

$$F(x) = \begin{cases} P(X \leq x) = \int_0^x \frac{1}{\theta} e^{-\frac{t}{\theta}}\, dt = 1 - e^{-\frac{x}{\theta}}, & x \geq 0; \\ 0, & \text{elsewhere.} \end{cases}$$

The exponential distribution function

Example

The life time, X, of a certain types of lighbulbs follows an exponential distribution with a mean of 100 hours. Plot the probability density function and the distribution function for X.

Solution

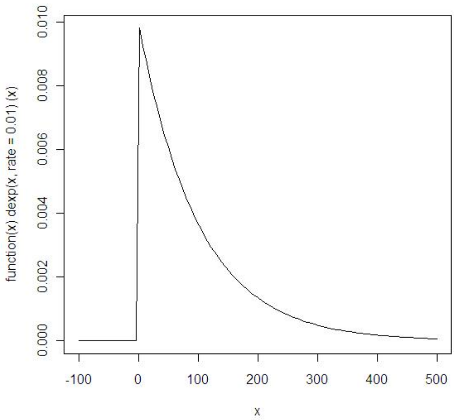
$\theta = 100.$

$$f(x) = \begin{cases} \frac{1}{100} e^{-\frac{x}{100}}, & x \geq 0; \\ 0, & \textit{elsewhere} \end{cases}$$

$$F(x) = \begin{cases} 1 - e^{-\frac{x}{100}}, & x \geq 0; \\ 0, & \textit{elsewhere.} \end{cases}$$

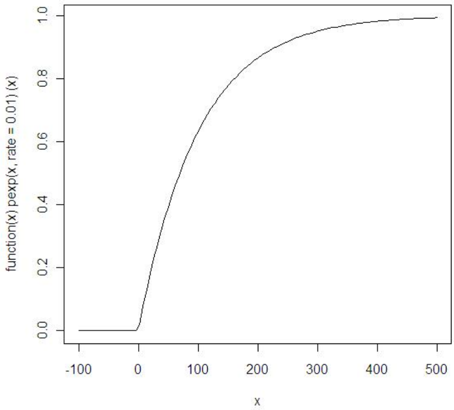
The exponential probability density function

> plot(function(x) dexp(x,rate=0.01), -100, 500)



The exponential distribution function

> plot(function(x) pexp(x,rate=0.01), -100, 500)



The Gamma function

Definition

The Gamma function is defined by

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x}\, dx$$

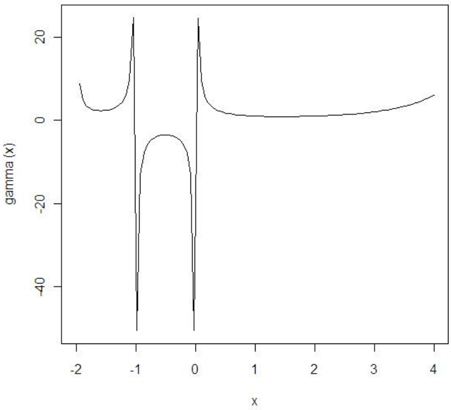
for $\alpha > 0$.

The Gamma function, Γ, has the following properties:

- $\Gamma(\alpha + 1) = \alpha\Gamma(\alpha)$
- $\Gamma(n) = (n-1)!$ for any integer $n \geq 1$.
- $\Gamma(n) = \infty$ for any integer $x \leq 0$.
- $\Gamma(\alpha)$ is infinitely differentiable.
- $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

Graph of the Gamma function

> plot(gamma,-2,4)



Let X be an exponential random variable with mean θ. Let, $y = \frac{1}{\theta}x$ such that $dy = \frac{1}{\theta}dx$. Then

$$\begin{aligned}
E(X^k) &= \frac{1}{\theta}\int_0^\infty x^k e^{-\frac{x}{\theta}}\,dx && (1)\\
&= \int_0^\infty (y\theta)^k e^{-y}\,dy && (2)\\
&= \theta^k \int_0^\infty y^k e^{-y}\,dy && (3)\\
&= \theta^k \Gamma(k+1) && (4)\\
&= \theta^k k!. && (5)
\end{aligned}$$

The expectations and the variance of the exponential distribution

- Letting $k = 1$ in (1), we obtain,

$$E(X) = \int_{-\infty}^{\infty} xf(x)\,dx = \int_{0}^{\infty} x\frac{1}{\theta}e^{-\frac{x}{\theta}}\,dx = \theta.$$

The expectations and the variance of the exponential distribution

- Letting $k = 1$ in (1), we obtain,

$$E(X) = \int_{-\infty}^{\infty} x f(x)\,dx = \int_{0}^{\infty} x \frac{1}{\theta} e^{-\frac{x}{\theta}}\,dx = \theta.$$

- Letting $k = 2$ in (1), we obtain,

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)\,dx = \int_{0}^{\infty} x^2 \frac{1}{\theta} e^{-\frac{x}{\theta}}\,dx = 2\theta^2.$$

The expectations and the variance of the exponential distribution

- Letting $k = 1$ in (1), we obtain,

$$E(X) = \int_{-\infty}^{\infty} xf(x)\,dx = \int_{0}^{\infty} x\frac{1}{\theta}e^{-\frac{x}{\theta}}\,dx = \theta.$$

- Letting $k = 2$ in (1), we obtain,

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x)\,dx = \int_{0}^{\infty} x^2\frac{1}{\theta}e^{-\frac{x}{\theta}}\,dx = 2\theta^2.$$

- Hence

$$V(X) = E(X^2) - \mu^2 = 2\theta^2 - \theta^2 = \theta^2.$$

The memoryless property

- Recall that the geometric distribution is the discrete distribution with the memoryless property.
- The exponential distribution is the continous distribution with the memoryless property:
- Suppse X is an exponential random variable with mean θ. Then

$$\begin{aligned}
P(X > a+b | X > a) &= \frac{P(X > a+b)}{P(X > a)} = \frac{1 - F(a+b)}{1 - F(a)} \\
&= \frac{1 - \left(1 - e^{-\left(\frac{a+b}{\theta}\right)}\right)}{1 - (1 - e^{-\frac{a}{\theta}})} \\
&= e^{-\frac{b}{\theta}} \\
&= 1 - F(b) = P(X > b).
\end{aligned}$$

The Poisson and the exponential distribution

- Suppose that the number of events, Y, in an hour period is following a Poisson distribution with mean λ, i.e. the rate is λ events per hour.
- Thus, in t hours, the number of events, Y, will have a Poisson distribution with mean value $\lambda \cdot t$.
- Let X denote the length of time until this first event. Then

$$P(X > t) = P[Y = 0 \text{ on the interval } (0, t)]$$
$$= \frac{(\lambda t)^0 e^{-\lambda t}}{0!} = e^{-\lambda t}$$

and

$$P(X \leq t) = 1 - P(X > t) = 1 - e^{-\lambda t}$$

which is the cumulative distribution function for the exponential distribution with $\lambda = \frac{1}{\theta}$.

The Poisson and the exponential distribution

- Thus, if the number of events in a specified interval has a Poisson distribution, the distance between any event and its successive event has an exponential distribution.

The Poisson and the exponential distribution

Example

Suppose the number of phone calls passing through a particular cellular relay system, follows a Poisson distribution with an average of 3 calls during a 2-min period.
(A) Find the probability, p, that no call will pass through the relay system during a given 2-min period.
(B) Find the probability that at least three minutes will pass before a call is passed through the relay system.

The Poisson and the exponential distribution

Solution

(A) Here $\lambda = 3$.
We have

$$p(0) = \frac{3^0}{0!} e^{-3} = 0.050.$$

(B) The mean number of phone calls per minute is $\frac{3}{2}$. Let Y be the length of time between two successive phone calls with a rate of $\frac{1}{\theta} = \frac{3}{2}$ per minute which follows an exponential distribution,

$$f(y) = \frac{3}{2} e^{-\frac{3}{2}y}, \quad y \geq 0.$$

The Poisson and the exponential distribution, solution cont.

Solution

$$P(Y \geq 3) = \frac{3}{2}\int_3^{\infty} e^{-\frac{3}{2}y}\,dy = e^{-\frac{9}{2}} = 0.011.$$

Notice that this is the same as

$$P(Y \geq 3) = P[X = 0 \text{ on the interval } (0,3)] = \frac{(\frac{3}{2}\cdot 3)^0 e^{-3\cdot\frac{3}{2}}}{0!} = e^{-\frac{9}{2}},$$

where X is the number of phone calls during a 1-minute period.

The exponential distribution

Example

Suppose the life length of a certain brand of automobiles follows an exponential distribution with a mean of 10000 miles.
(A) Find the probability that one of these tires, bought today, will last over 10000 miles.
(B) Find the probability that at least one of the four tires bought today will last over 10000 miles.

The exponential distribution

Solution

$\theta = 10000$. *Let* X *be the number of miles a tire last.*

(A)

$$P(X \geq 10000) = \frac{1}{10000}\int_{10000}^{\infty} e^{-\frac{x}{10000}}\,dx = e^{-1} = 0.368.$$

(B) Let Y *be the number of tires that last more than 10000 miles.* Y *follows a Binomial distribution with probability of success,* $p = 0.368$. *We have*

$$P(Y \geq 1) = 1 - P(Y = 0) = 1 - (1 - 0.368)^4 = 0.840.$$

The Gamma distribution

Distributions that have low probabilites for intervals close to zero, with increasing probability up to a point followed by a decreasing probability.

Theorem

The gamma distribution:

$$f(x) = \begin{cases} \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} e^{-\frac{x}{\beta}}, & x \geq 0; \\ 0, & \textit{elsewhere} \end{cases}$$

$$E(X) = \alpha\beta \;\; \textit{and} \;\; V(X) = \alpha\beta^2, \;\; \textit{for} \;\; \alpha, \beta > 0.$$

The parameters α and β determines the shape of the curve.

The Gamma distribution

Recall from (1) that

$$\int_0^\infty x^\alpha e^{-\frac{x}{\beta}}\, dx = \beta^{\alpha+1}\Gamma(\alpha+1).$$

We have

$$\begin{aligned}\int_{-\infty}^\infty f(x)dx &= \int_0^\infty \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-\frac{x}{\beta}}\, dx \\ &= \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty x^{\alpha-1} e^{-\frac{x}{\beta}}\, dx \\ &= \frac{1}{\Gamma(\alpha)\beta^\alpha} \beta^\alpha \Gamma(\alpha) = 1.\end{aligned}$$

The expectation and the variance of the gamma distribution

$$
\begin{aligned}
E(X) &= \int_0^\infty x \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-\frac{x}{\beta}} dx \\
&= \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty x^\alpha e^{-\frac{x}{\beta}} dx \\
&= \frac{1}{\Gamma(\alpha)\beta^\alpha} \Gamma(\alpha+1)\beta^{\alpha+1} = \alpha\beta.
\end{aligned}
$$

Similarly, we have

$$E(X^2) = \alpha(\alpha+1)\beta^2$$

and hence

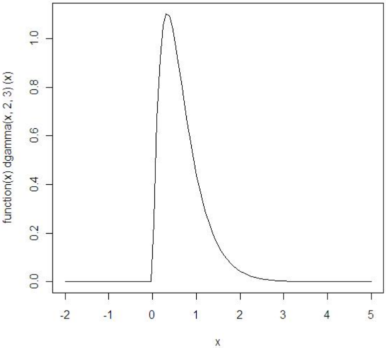
$$V(X) = E(X^2) - \mu^2 = \alpha(\alpha+1)\beta^2 - \alpha^2\beta^2 = \alpha\beta^2.$$

Graph of the gamma probability density function

$>$ plot(function(x) dgamma(x,α,β))

For $\alpha = 2$ and $\beta = 3$, we have:

$>$ plot(function(x) dgamma(x,2,3), -2, 5)



The gamma and the exponential distribution

- Notice that when $\alpha = 1$, the gamma density function is reduced to the exponential density function.
- If X_1, X_2, ..., X_n represents identically distributed, independent gamma random variables with parameters α and β, then

$$Y = \sum_{i=1}^{n} X_i$$

also have a gamma distribution with parameters $n\alpha$ and β. Hence

$$E(Y) = n\alpha\beta \text{ and } V(Y) = n\alpha\beta^2.$$

The gamma and the exponential distribution

- Suppose we have light bulbs.
- Suppose the time that each bulb will burn is exponentially distributed with rate β.
- Suppose the life length of one bulb is independent of others.
- The time until the α^{th} ceases to burn follows a gamma distribution with parameters α and β.

The gamma and the exponential distribution

Example

Three lifebulbs has lifelengths of X_1, X_2, and X_3 respectively in which all follows an exponential distribution with a mean of 200 hours. Suppose the life length of one bulb is independent of others. Find the probability distribution and expected value for the time until all three bulbs ceases to burn.

The gamma and the exponential distribution

Solution

Let $Y = X_1 + X_2 + X_3$ *denote the total life length.* Y *has a gamma distribtution with parameters* $\alpha = 3$ *and* $\beta = 200$*. Thus*

$$f(y) = \begin{cases} \frac{1}{\Gamma(3)(200)^3} x e^{-\frac{x}{200}}, & x \geq 0; \\ 0, & elsewhere \end{cases}$$

$$E(Y) = \alpha\beta = 3 \cdot 200 = 600 \textit{ hours.}$$

The Normal distribution

- The normal distribution is the most frequently occuring continuous probability distributions.
- Much statistics theory is based on the normal distribution.
- Central limit theorem

Examples:

- Heights of adults
- Observations errors in an experiment
- Velocities of the molecules in the ideal gas.
- Many naturally occuring measurements have relative frequency distributions that closely fit the normal curve.

The Normal distribution

Theorem

The Normal distribution:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \;\; -\infty < x < \infty$$

$$E(X) = \mu \;\; and \;\; V(X) = \sigma^2.$$

The Normal distribution

Theorem

$$\int_{-\infty}^{\infty} e^{-x^2}\, dx = \sqrt{\pi}.$$

Theorem

If X has a normal distribution with mean μ and a variance of σ^2, and

$$Y = aX + b,$$

where a, b are constants, then Y also has a normal distribution with

$$E(Y) = a\mu + b \;\text{ and }\; V(Y) = a^2\sigma^2.$$

The Standard Normal distribution

Definition

We say that Z has a standard normal distribution if it has a normal distribution with $\mu = 0$ and $\sigma = 1$.

Define

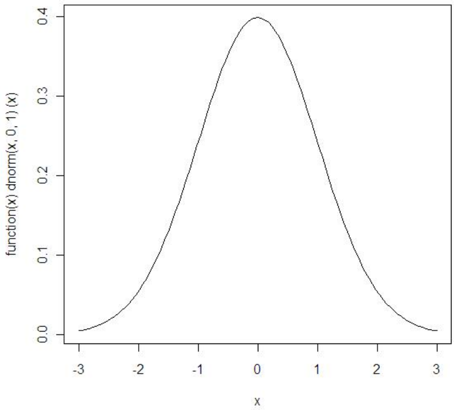
$$Z = \frac{X - \mu}{\sigma},$$

where X is a random variable that have a normal distribution with mean μ and standard deviation σ. Then Z has a standard normal distribution.
Tables of integrals, back in the book, gives numerically values for

$$P(0 \leq Z \leq z) = \frac{1}{\sqrt{2\pi}} \int_0^z e^{-\frac{t^2}{2}}\, dt.$$

The Standard Normal distribution

> plot(function(x) dnorm(x,0,1),-3,3)



The Normal distribution

- The probability density function, $f(x)$, of a normal random variable with a mean of μ and a standard deviation of σ has points of inflection at $x = \mu \pm \sigma$. The proof of this is left as a homework problem.

Example

Let Z denote a standard normal variable. Find
(A) $P(Z \leq 1)$
(B) $P(Z \geq 1)$
(C) $P(Z < -1.5)$
(D) $P(-1 \leq Z \leq 1.5)$
(E) Find a value z_0 such that $P(Z \leq z_0) = 0.85$.

The Normal distribution

Solution

(A)

$$P(Z \leq 1) = P(Z \leq 0) + P(0 \leq Z \leq 1) = 0.5 + 0.3413 = 0.8413.$$

(B)

$$P(Z \geq 1) = 1 - P(Z \leq 1) = 1 - 0.8413 = 0.1587.$$

(C)

$$\begin{aligned} P(Z < -1.5) &= P(Z > 1.5) \\ &= 1 - (0.5 + P(0 \leq Z \leq 1.5)) \\ &= 0.5 - 0.4332 = 0.0668. \end{aligned}$$

The Normal distribution, Solution cont.

Solution

(D)

$$\begin{aligned} P(-1 \leq Z \leq 1.5) &= P(-1 \leq Z \leq 0) + P(0 \leq Z \leq 1.5) \\ &= P(0 \leq Z \leq 1) + P(0 \leq Z \leq 1.5) \\ &= 0.3413 + 0.4332 = 0.7745. \end{aligned}$$

The Normal distribution, Solution cont.

Solution

(E) By table:

$$P(Z \leq z_0) = 0.5 + P(0 \leq Z \leq z_0) = 0.85.$$

Hence

$$P(0 \leq Z \leq z_0) = 0.35.$$

From the table, we find $z_0 = 1.04$.
By using R, we find,
$> z_0 = qnorm(0.85)$
[1] 1.0364.

The Standard Normal distribution, the R-code

If $P(Z \le z_0) = p$, then in R,
$> z_0 = qnorm(p)$.
To compute $F(z_0) = \int_{-\infty}^{z_0} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} dz$, in R:
$> pnorm(z_0)$.
Example:
$> pnorm(2)$
[1] 0.9772.

The Standard Normal distribution

We have

$$P(-1 \leq Z \leq 1) = 2P(0 \leq Z \leq 1) = 2 \cdot 0.3413 = 0.683.$$

$$P(-2 \leq Z \leq 2) = 2 \cdot 0.4772 = 0.954.$$

$$P(-3 \leq Z \leq 3) = 2 \cdot 0.4987 = 0.997.$$

Thus, for a standard normal distribution, approximately,
68% of the values fall within 1 standard deviation of the mean in either direction.
95% of the values fall within 2 standard deviation of the mean in either direction.
99.7% of the values fall within 3 standard deviation of the mean in either direction.

The Normal distribution

Example

- Find $P(4 \leq X \leq 12)$, where X is a normal random variable that has a mean of 8 and a standard deviation of 2.

The Normal distribution

Example

- Find $P(4 \leq X \leq 12)$, where X is a normal random variable that has a mean of 8 and a standard deviation of 2.

Solution

Let

$$Z = \frac{X-8}{2}.$$

Then

$$P(4 \leq X \leq 12) = P\left(\frac{4-8}{2} \leq Z \leq \frac{12-8}{2}\right)$$
$$= P(-2 \leq Z \leq 2) = 0.954.$$

The Normal distribution, R-code

In R, the problem can also be solved in this way:

$$P(4 \leq X \leq 12) = P(X \leq 12) - P(X \leq 4).$$

Then type
$>$ pnorm(12,8,2)-pnorm(4,8,2)
[1] 0.954.
In general, in R:
$> pnorm(x, \mu, \sigma)$
which is the cumulative normal distribution, $P(X \leq x)$, with mean μ and standard deviation σ.

Example

If X has a normal distribution with mean 2 and standard deviation 2, find $P(X^2 - 4X \leq 12)$.

Example

If X has a normal distribution with mean 2 and standard deviation 2, find $P(X^2 - 4X \leq 12)$.

Solution

$$\begin{aligned} P(X^2 - 4X \leq 12) &= P(X^2 - 4X + 4 \leq 16) \\ &= P((X-2)^2 \leq 16) \\ &= P(-4 \leq (X-2) \leq 4) \\ &= P(-2 \leq \frac{X-2}{2} \leq 2) \\ &= P(-2 \leq Z \leq 2) = 0.95, \end{aligned}$$

where $Z = \frac{X-2}{2}$.