ST 501: Intro Probability & Distribution Theory

Continuous Random Variables (part II)

Fall 2024

Department of Statistics, North Carolina State University.

Gamma function

We now introduce the Γ function, arguably the most common of the special functions or the most special of the common functions.

Definition
Let $z > 0$ and define $\Gamma(z)$ via the definite integral

$$\Gamma(z) = \int_0^\infty x^{z-1} e^{-x} \, \mathrm{d}x$$

Coffee break Verify, at your leisure, that $\Gamma(z)$ is well-defined for $z > 0$.

A key property of $\Gamma(z)$ is the following recurrence formula

Proposition
Let $z > 0$. Then

$$\Gamma(z+1) = z\Gamma(z).$$

A key property of $\Gamma(z)$ is the following recurrence formula

Proposition
Let $z > 0$. Then

$$\Gamma(z+1) = z\Gamma(z).$$

Proof This is an exercise in integration by parts.

Let $u = x^z$ and $\mathrm{d}v = e^{-x}\mathrm{d}x$.

Then $\mathrm{d}u = zx^{z-1}\,\mathrm{d}x$ and $v = -e^{-x}$. Hence

$$\begin{aligned}\Gamma(z+1) &= \int_0^\infty x^z e^{-x}\,\mathrm{d}x = -x^z e^{-x}\Big|_0^\infty + \int_0^\infty zx^{z-1}e^{-x}\,\mathrm{d}x \\ &= 0 + z\Gamma(z) = z\Gamma(z).\end{aligned}$$

3

We note that $\Gamma(1) = 1$. Therefore, if z is a positive integer,

$$\begin{aligned}\Gamma(z+1) &= z\Gamma(z)\\ &= z(z-1)\Gamma(z-1)\\ &= z(z-1)(z-2)\Gamma(z-2) = \cdots = z!\end{aligned}$$

$\Gamma(z)$ is thus an extension of the factorial function from the positive integers to the positive real numbers.

$\Gamma(1/2) = \sqrt{\pi}$ (optional)

Let $x = u^2$. Then $\mathrm{d}x = 2u\mathrm{d}u$ and hence

$$\Gamma(z) = \int_0^\infty x^{z-1} e^{-x}\, \mathrm{d}x = \int_0^\infty u^{2(z-1)} e^{-u^2}\, 2u\, \mathrm{d}u.$$

In other words

$$\Gamma(1/2) = 2\int_0^\infty e^{-u^2}\, \mathrm{d}u.$$

Recall the integral identity for the pdf of the normal distribution. Then

$$\Gamma(1/2) = 2\int_0^\infty e^{-u^2}\, \mathrm{d}u = \int_{-\infty}^\infty e^{-u^2}\, \mathrm{d}u = \sqrt{\pi}.$$

5

Gamma distributions

Definition
Let $\alpha > 0$ and $\beta > 0$. X is a said to be a gamma rv with shape parameter α and scale parameter β if X has pdf

$$f(x) = \frac{x^{\alpha-1}e^{-x/\beta}}{\Gamma(\alpha)\beta^{\alpha}}, \quad x \geq 0.$$

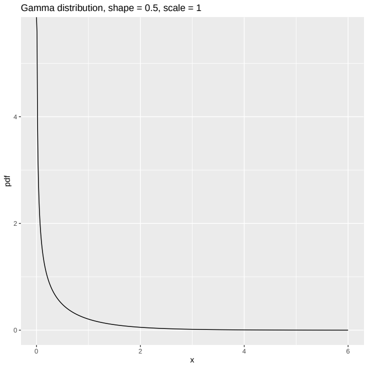
Equivalently, X is said to be a gamma rv with shape parameter α and rate parameter $\lambda = 1/\beta$ if X has pdf

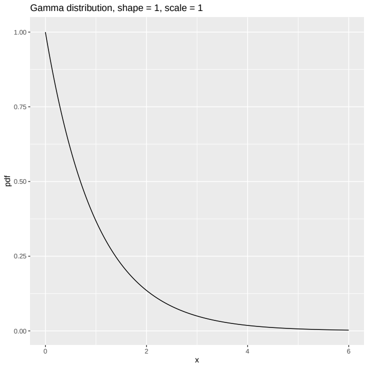
$$f(x) = \frac{\lambda^{\alpha}x^{\alpha-1}e^{-\alpha x}}{\Gamma(\alpha)}, \qquad x \geq 0$$

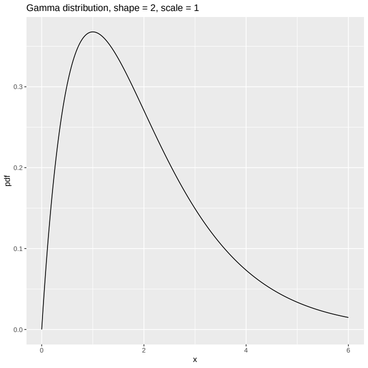
We write $X \sim \Gamma(\alpha, \beta)$ or $X \sim \Gamma(\alpha, \lambda)$ to denote a gamma rv with shape parameter α and scale parameter β (or rate parameter $\lambda = 1/\beta$).

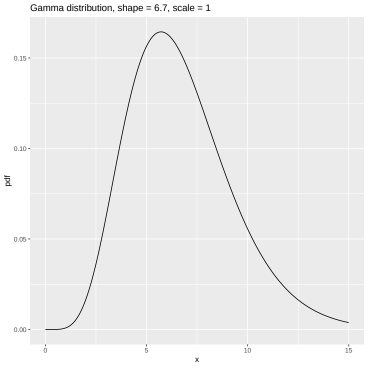
Shape parameter of Gamma distributions

As α changes, the shape of the pdf changes.



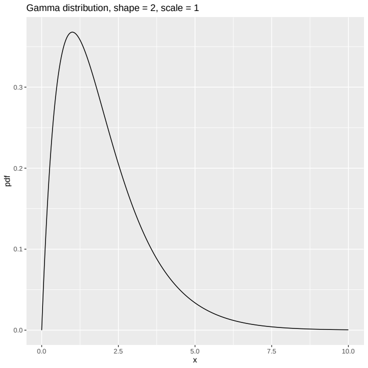


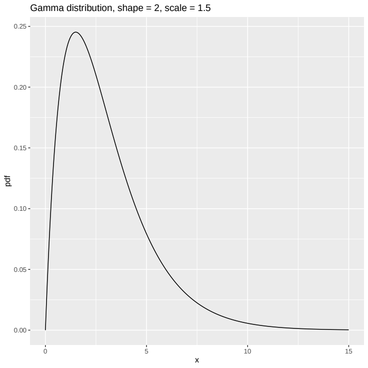

Gamma distribution, shape = 2, scale = 1
pdf
0.0
0.1
0.2
0.3
0
2
4
6
x

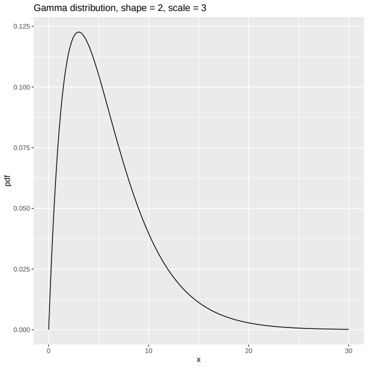

Gamma distribution, shape = 6.7, scale = 1
pdf
0.00
0.05
0.10
0.15
0
5
10
15
x

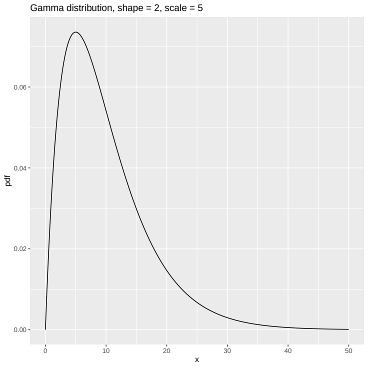
Scale parameter of Gamma distributions

As β changes, only the scale changes.









Proposition

let $X \sim \Gamma(\alpha, \beta)$ be a gamma rv with shape parameter α and scale parameter β. Then for any $c > 0$, $cX \sim \Gamma(\alpha, c\beta)$.

10

Exponential distribution

Definition
Let $X \sim \text{Gamma}(1, \beta)$ where β is the scale parameter. Then X is said to be an exponential random variable with rate $\lambda = 1/\beta$. We denote this as $X \sim \text{Exp}(\lambda)$.

If $X \sim \text{Exp}(\lambda)$ then X has pdf

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

11

The cdf of $X \sim \mathrm{Exp}(\lambda)$ is

$$F(x) = P(X \leq x) = \int_0^x \lambda e^{-\lambda t}\,\mathrm{d}t = -e^{-\lambda t}\Big|_0^x = 1 - e^{-\lambda x}$$

The tail probability for $X \sim \mathrm{Exp}(\lambda)$ is

$$P(X \geq x) = 1 - F(x) = e^{-\lambda x}.$$

and the quantile function F^{-1} is

$$F^{-1}(p) = -\frac{\ln(1-p)}{\lambda}, \qquad 0 < p < 1$$

12

From Poisson to Exponential

We are interested in modeling X, the waiting time until the first occurrence of a particular event. Suppose furthermore that, on average, λ of these particular events happen during one unit of time. What is a reasonable distribution for X ?

13

From Poisson to Exponential

We are interested in modeling X, the waiting time until the first occurrence of a particular event. Suppose furthermore that, on average, λ of these particular events happen during one unit of time. What is a reasonable distribution for X ?

A. Let $x > 0$ be given. A plausible model for Z_x, the number of events appearing during the time interval $[0, x]$, is $Z_x \sim \mathrm{Pois}(\lambda x)$. Now $X \geq x$ if and only if $Z_x = 0$. We thus have

$$P(X \geq x) = P(Z_x = 0) = e^{-\lambda x} \Longrightarrow P(X \leq x) = 1 - e^{-\lambda x}$$

which is the cdf for an $\mathrm{Exp}(\lambda)$ random variable.

Note Exponential rv is commonly used to model the waiting time until the first failure of some equipment, machines, etc. The parameter λ is then the "failure rate", with larger λ corresponding to more frequent failures.

Exponential distribution is memoryless

Proposition

Let $X \sim \mathrm{Exp}(\lambda)$. *Then* X *is memoryless, i.e., for all* $t > 0$, $s > 0$

$$P(X \geq t + s \mid X \geq t) = P(X \geq s) = e^{-\lambda s}.$$

14

Exponential distribution is memoryless

Proposition
Let $X \sim \mathrm{Exp}(\lambda)$. *Then* X *is memoryless, i.e., for all* $t > 0$, $s > 0$

$$P(X \geq t + s \mid X \geq t) = P(X \geq s) = e^{-\lambda s}.$$

Proof Direct calculations yield

$$P(X \geq t + s \mid X \geq t) = \frac{P(X \geq t + s)}{P(X \geq t)} = \frac{e^{-\lambda(t+s)}}{e^{-\lambda t}} = P(X \geq s)$$

Note The exponential rv is the unique continuous rv with the memoryless property.

14

Sum of exponential rv (optional)

Proposition

Let $X_1, X_2, \ldots, X_k$ *be* k *mutually independent* $\mathrm{Exp}(\lambda)$ *random variables. Then* $X = X_1 + X_2 + \cdots + X_m$ *is a gamma rv with shape parameter* $\alpha = m$ *and rate parameter* λ *(equivalently, scale parameter* $\beta = 1/\lambda$*.*

15

Proof Since each of the X_i is the waiting time until the first occurrence of an event, we can view $X = X_1 + X_2 + \cdots + X_m$ as the waiting time until the mth event. Now consider $P(X \geq x)$ for a given $x > 0$.

Let Z_x be the Poisson rv for the number of events during the time interval $[0, x]$. Then $Z_x \sim \mathrm{Pois}(\lambda x)$. Now $X \geq x$ if and only if $Z_x \leq m - 1$. We therefore have

$$P(X \leq x) = 1 - P(Z_x \leq m-1) = 1 - \sum_{i=0}^{m-1} \frac{e^{-\lambda x}(\lambda x)^i}{i!}$$

16

Taking derivative of $P(X \leq x)$ wrt x, the pdf for X is then

$$
\begin{aligned}
f(x) = \frac{\mathrm{d}}{\mathrm{d}x} P(X \leq x) &= \sum_{i=0}^{m-1} \frac{\mathrm{d}}{\mathrm{d}x} \frac{-e^{-\lambda x}(\lambda x)^i}{i!} \\
&= \sum_{i=0}^{m-1} \frac{\lambda e^{-\lambda x}(\lambda x)^i - i\lambda e^{-\lambda x}(\lambda x)^{i-1}}{i!} \\
&= \sum_{i=0}^{m-1} \frac{e^{-\lambda x}\lambda^{i+1}x^i}{i!} - \sum_{i=0}^{m-1} \frac{ie^{-\lambda x}\lambda^i x^{i-1}}{i!} \\
&= \sum_{i=0}^{m-1} \frac{e^{-\lambda x}\lambda^{i+1}x^i}{i!} - \sum_{i=0}^{m-2} \frac{e^{-\lambda x}\lambda^{i+1}x^i}{i!} \\
&= \frac{\lambda^m x^{m-1} e^{-\lambda x}}{(m-1)!}
\end{aligned}
$$

which is the pdf of a $\Gamma(\alpha = m, \beta = 1/\lambda)$.

17

A gamma rv with shape parameter α where α is a positive integer is known as an Erlang random variable. An Erlang rv is thus a sum of independent exponential rv. The exponential and the Erlang distribution is the continuous analog of the geometric and the negative binomial distribution, respectively.

Note In general there is no ”simple formula” for the cdf of a gamma rv except in the case when α is a positive integer.

18

68. The special case of the gamma distribution in which α is a positive integer n is called an Erlang distribution. If we replace β by $1/\lambda$ in Expression (4.8), the Erlang pdf is

$$f(x; \lambda, n) = \begin{cases} \dfrac{\lambda(\lambda x)^{n-1}e^{-\lambda x}}{(n-1)!} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

It can be shown that if the times between successive events are independent, each with an exponential distribution with parameter λ, then the total time X that

19

Example The lifetime (in hours) X of an electronic component is a random variable with pdf

$$f(x) = \frac{1}{100} e^{-x/100}, \qquad x > 0$$

Three of these components operate independent in a piece of equipment. The equipment fails if at least two of the components fail. Find the probability that the equipment will operate for at least 200 hours without failure.

20

We see that $X \sim \mathrm{Exp}(1/100)$, i.e., the rate of failure is 1/100 every hour, or once every 100 hours.

Let Z be the waiting time until the second failure.

We are interested in $P(Z > 200)$. Here Z is an Erlang rv with shape parameter 2 and rate parameter $\lambda = 1/100$.

Now let $W \sim \mathrm{Pois}(\lambda \times 200)$ be the rv modeling the number of failures during time $[0, 200]$. Then

$$P(Z > 200) = P(W \le 1) = e^{-2} + e^{-2} \times 2 = 3e^{-2} \approx 0.406$$

21

Chi-square distributions

Definition
Let $m > 0$ and let $X \sim \Gamma(m/2, 2)$, i.e., X is a gamma rv with shape parameter $\alpha = m/2$ and scale parameter $\beta = 2$. Then X is said to be a chi-squared rv with m degree(s) of freedom. The pdf for X is

$$f(x) = \frac{x^{m/2-1}e^{-x/2}}{\Gamma(m/2)2^{m/2}}, \quad x > 0.$$

We denote a chi-squared rv with m degrees of freedom as $X \sim \chi^2_m$. m is usally a positive integer but this is not required.

22

When $m = 1$, the pdf of $X \sim \chi_1^2$ is (recall that $\Gamma(1/2) = \sqrt{\pi}$)

$$f(x) = \frac{e^{-x/2}}{\sqrt{2\pi x}}, \qquad x > 0$$

When $m = 2$, the pdf of $X \sim \chi_2^2$ is

$$f(x) = \frac{1}{2} e^{-x/2}$$

which is an exponential rv with rate parameter $\lambda = 1/2$.

Proposition

If $Z \sim \mathcal{N}(0, 1)$ is standard normal then $X = Z^2$ is chi-squared with $m = 1$ degree of freedom.

23

Proof Since Z take on both positive and negative values, Z^2 is not monotone and so we cannot directly apply the result for the monotone transformation of a pdf.

Nevertheless, Z^2 is monotone wrt $|Z|$ and hence

$$P(X \leq x) = P(Z^2 \leq x) = P(-\sqrt{x} \leq Z \leq \sqrt{x}) = \Phi(\sqrt{x}) - \Phi(-\sqrt{x})$$

Taking derivative of $P(X \leq x)$ wrt x, we obtain

$$f(x) = \frac{\mathrm{d}}{\mathrm{d}x}\big(\Phi(\sqrt{x}) - \Phi(-\sqrt{x})\big) = \frac{e^{-x/2}}{\sqrt{2\pi x}}.$$

24

Mean and variance of gamma rv

Proposition

Let $X \sim \Gamma(\alpha, \beta)$ where α and β are the shape and scale parameter. Then

$$\mathbb{E}[X] = \alpha\beta = \frac{\alpha}{\lambda}, \quad \mathrm{Var}[X] = \alpha\beta^2 = \frac{\alpha}{\lambda^2},$$

where $\lambda = 1/\beta$ is the rate parameter.

25

Proof Direct calculations yield

$$\mathbb{E}[X] = \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty x^\alpha e^{-x/\beta}\,\mathrm{d}x = \frac{\Gamma(\alpha+1)\beta^{\alpha+1}}{\Gamma(\alpha)\beta^\alpha} = \alpha\beta$$

as $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$. Similarly,

$$\mathbb{E}[X^2] = \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty x^{\alpha+1} e^{-x/\beta}\,\mathrm{d}x = \frac{\Gamma(\alpha+2)\beta^{\alpha+2}}{\Gamma(\alpha)\beta^\alpha} = \alpha(\alpha+1)\beta^2$$

and so $\mathrm{Var}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \alpha\beta^2$.

26

Example Let $X \sim \mathrm{Exp}(\lambda)$ where λ is the rate parameter. Then $\mathbb{E}[X] = \lambda^{-1}$ and $\mathrm{Var}[X] = \lambda^{-2}$.

Example Let $X \sim \chi^2_m$ be a chi-squared r.v. with m degrees of freedom. Then X is a gamma r.v. with shape parameter $\alpha = m/2$ and scale parameter $\beta = 2$. Hence $\mathbb{E}[X] = m$ and $\mathrm{Var}[X] = 2m$.

27

Higher order moments as well as inverse and fractional moments of a gamma rv are also easy to compute.

Proposition
Let $X \sim \Gamma(\alpha, \beta)$*. Then for all* $c > -\alpha$*,*

$$\mathbb{E}[X^c] = \frac{1}{\Gamma(\alpha)\beta^\alpha} \int_0^\infty x^{c+\alpha-1} e^{-x/\beta} \, \mathrm{d}x = \frac{\Gamma(c+\alpha)\beta^c}{\Gamma(\alpha)}$$

28

mgf of gamma rv

Theorem
Let $X \sim \Gamma(\alpha, \beta)$ *where* β *is the scale parameter. Then*

$$m_X(t) = \left(\frac{1}{1-\beta t}\right)^{\alpha}, \qquad t \in \mathbb{R},\ t < 1/\beta$$

29

Proof We recall that the pdf f of a $\Gamma(\alpha, \beta)$ r.v. implies

$$f(x) = \frac{x^{\alpha-1}e^{-x/\beta}}{\Gamma(\alpha)\beta^\alpha} \implies \int_0^\infty x^{\alpha-1}e^{-x/\beta}\,\mathrm{d}x = \Gamma(\alpha)\beta^\alpha.$$

The result follows directly from the above display, i.e.,

$$\begin{aligned} m_X(t) = \mathbb{E}[e^{tX}] &= \frac{1}{\Gamma(\alpha)\beta^\alpha}\int_0^\infty e^{tx}x^{\alpha-1}e^{-x/\beta}\,\mathrm{d}x \\ &= \frac{1}{\Gamma(\alpha)\beta^\alpha}\int_0^\infty x^{\alpha-1}e^{-(1/\beta-t)x}\,\mathrm{d}x \\ &= \frac{1}{\Gamma(\alpha)\beta^\alpha}\int_0^\infty x^{\alpha-1}e^{-(1-\beta t)x/\beta}\,\mathrm{d}x \\ &= \frac{\Gamma(\alpha)(\beta/(1-\beta t))^\alpha}{\Gamma(\alpha)\beta^\alpha} = \left(\frac{1}{1-\beta t}\right)^\alpha. \end{aligned}$$

Important For the above calculations to be valid, we need

$$\int_0^\infty x^{\alpha-1} e^{-(1/\beta - t)x} \, \mathrm{d}x$$

to be well-defined, i.e., finite.

This only happens when $1/\beta - t > 0$, i.e., we require $t < 1/\beta$ as in the statement of the theorem.

31

The form of the mgf for $\Gamma(\alpha, \beta)$ implies the following immediate corollary.

Corollary
Let $X_1, X_2, \ldots, X_m$ be mutually independent random variables with $X_i \sim \Gamma(\alpha_i, \beta)$, i.e., the X_i have the same scale parameter β. Then $X = X_1 + X_2 + \cdots + X_m$ is a $\Gamma(\sum_i \alpha_i, \beta)$.

That is to say, the sum of independent gamma r.v.s (with a common scale parameter) also has a gamma distribution.

32

Example Since a chi-squared r.v. with 1 degree of freedom is the square of a standard normal, the above corollary implies that a chi-square with m degrees of freedom (for integer $m \geq 1$) is the sum of squares of m mutually independent standard normal, i.e,

$$X \sim \chi^2_m \Longleftrightarrow X = Z_1^2 + Z_2^2 + \cdots + Z_m^2$$

where $Z_1, Z_2, \ldots, Z_m$ are mutually independent $\mathcal{N}(0, 1)$ r.v.s. Thus $X \sim \chi^2_m$ is a $\Gamma(m/2, 2)$ r.v. with mgf

$$m_X(t) = \left(\frac{1}{1-2t}\right)^{m/2} = (1-2t)^{-m/2}, \qquad t < 1/2.$$

33

Weibull distribution

Definition
Let $\alpha > 0$ and $\beta > 0$. A rv X is said to have a Weibull distribution with shape parameters α and β if X has pdf

$$f(x) = \frac{\alpha}{\beta^{\alpha}} x^{\alpha-1} \exp(-(x/\beta)^{\alpha}), \quad x \geq 0.$$

We write $X \sim \mathrm{Weibull}(\alpha, \beta)$. See problem 2.67 in your textbok.

34

From Exponential to Weibull

Proposition

Let $Y \sim \mathrm{Exp}(1)$. *Then* $\beta Y^{1/\alpha} \sim \mathrm{Weibull}(\alpha, \beta)$.

35

Proposition
Let $Y \sim \mathrm{Exp}(1)$*. Then* $\beta Y^{1/\alpha} \sim \mathrm{Weibull}(\alpha, \beta)$.

Proof Let $X = \beta Y^{1/\alpha}$. Then

$$P(X \leq x) = P(Y \leq (x/\beta)^\alpha) = 1 - \exp(-(x/\beta)^\alpha) \quad (*)$$

Taking the derivative of $P(X \leq x)$ wrt x yields

$$f(x) = \frac{\mathrm{d}}{\mathrm{d}x}\Big(1 - \exp(-(x/\beta)^\alpha)\Big) = \frac{\alpha}{\beta^\alpha} x^{\alpha-1} \exp(-(x/\beta)^\alpha).$$

35

Quantiles The cdf of a $\mathrm{Weibull}(\alpha, \beta)$ is given in Eq. (*). In particular the quantile function for Weibull is

$$1 - \exp(-(x/\beta)^{\alpha}) = p \Longrightarrow F^{-1}(p) = \beta\left(\log \frac{1}{1-p}\right)^{1/\alpha}.$$

36

Mean and variance for Weibull

Proposition

Let $X \sim \mathrm{Weibull}(\alpha, \beta)$. *Then*

$$\mathbb{E}[X] = \beta\Gamma\Big(1 + \frac{1}{\alpha}\Big), \quad \mathrm{Var}[X] = \beta^2\Big[\Gamma\Big(1 + \frac{2}{\alpha}\Big) - \Gamma\Big(1 + \frac{1}{\alpha}\Big)^2\Big].$$

37

Proof Change of variable together with the Γ integral.

Let $z = (x/\beta)^\alpha$. Then $\mathrm{d}z = \alpha\beta^{-\alpha}x^{\alpha-1}\,\mathrm{d}x$. Thus, for any $c > -\alpha$,

$$\begin{aligned}\mathbb{E}[X^c] &= \int_0^\infty \frac{\alpha}{\beta^\alpha} x^{\alpha+c-1} e^{-(x/\beta)^\alpha}\,\mathrm{d}x \\ &= \int_0^\infty \beta^c z^{c/\alpha} e^{-z}\,\mathrm{d}z \\ &= \beta^c \Gamma\left(1 + \frac{c}{\alpha}\right).\end{aligned}$$

Substituting $c = 1$ and $c = 2$ yields

$$\mathbb{E}[X] = \beta\Gamma\left(1 + \frac{1}{\alpha}\right), \quad \mathbb{E}[X^2] = \beta^2\Gamma\left(1 + \frac{2}{\alpha}\right).$$

38

72. The lifetime X (in hundreds of hours) of a certain type of vacuum tube has a Weibull distribution with parameters $\alpha = 2$ and $\beta = 3$. Compute the following:

a. $E(X)$ and $V(X)$

b. $P(X \leq 6)$

c. $P(1.5 \leq X \leq 6)$

(This Weibull distribution is suggested as a model for time in service in **"On the Assessment of Equipment Reliability: Trading Data Collection Costs for Precision,"** ***J. of Engr. Manuf.*, 1991: 105–109.)**

Recall that $\Gamma(1/2) = \sqrt{\pi}$ and $\Gamma(3/2) = 1/2\Gamma(1/2)$. Thus

$$\mathbb{E}[X] = 3\Gamma(3/2) \approx 2.659, \quad \mathrm{Var}[X] = 9\Gamma(2) - 9\Gamma(3/2)^2 \approx 1.931.$$

The cdf of X is $F(x) = 1 - \exp(-(x/3)^2)$ and hence

$$F(6) = 1 - \exp(-4) \approx 0.982, \quad F(1.5) = 1 - \exp(-1/4) \approx 0.221.$$

Beta function

Another special function commonly use in probability.

Definition

Let $\alpha > 0$ and $\beta > 0$. The beta function with parameters α and β is defined via the definite integral

$$B(\alpha, \beta) = \int_0^1 t^{\alpha-1}(1-t)^{\beta-1}\,\mathrm{d}t$$

The following result links the beta and Γ functions.

Proposition
For any $\alpha > 0$ *and* $\beta > 0$ *we have*

$$B(\alpha, \beta) = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}.$$

41

Beta distribution

Definition
A rv X is said to have a beta distribution with parameters $\alpha > 0$ and $\beta > 0$ if X has pdf

$$f(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)}, \quad 0 < x < 1.$$

We write $X \sim \mathrm{Beta}(\alpha, \beta)$ for a beta rv with parameters α and β.

Try this Go to this link and play around with the α and β parameters (you can input your own numbers) to see the diverse shapes of Beta rv.

42

Example When $\alpha = \beta = 1$ then $B(\alpha, \beta) = 1$ and $\mathrm{Beta}(\alpha, \beta)$ reduces to $\mathrm{Unif}(0, 1)$, the uniform distribution on $[0, 1]$.

Example Let $U \sim \mathrm{Unif}(0, 1)$. Recall, for $\alpha > 0$, $X = U^\alpha$ has pdf

$$f(x) = \frac{1}{\alpha} x^{1/\alpha - 1}$$

which coincides with the pdf of a $\mathrm{Beta}(1/\alpha, 1)$ random variable.

Example Let $X \sim \mathrm{Beta}(\alpha, \beta)$. Then $1 - X \sim \mathrm{Beta}(\beta, \alpha)$.

43

Example The pdf of a $X \sim \mathrm{Beta}(\alpha, \beta)$ rv is non-zero on the interval $[0, 1]$.

By shifting and scaling X via $(B - A)X + A$ we get a "beta-like" rv on the interval $[A, B]$ for $B > A$

In particular $Z = (B - A)X + A$ has pdf

$$f(z) = \frac{(z - A)^{\alpha-1}(B - z)^{\beta-1}}{(B - A)^{\alpha+\beta-1}\mathrm{Beta}(\alpha, \beta)}.$$

Distribution over distributions (optional)

The beta distribution is known as a ”probability distribution over probability distributions”.

In simple words, a beta rv take values in $[0, 1]$ and many discrete distributions (binomial, geometric, negative binomial, hypergeometric) are constructed using Bernoulli trials.

If we assume that p (the probability of success of the Bernoulli trials) are ”unknown”, then a plausible approach is to assume that p is a ”random variable” taking values in $[0, 1]$.

We then model p as a beta rv, and by choosing the parameters α and β, allow us to generate a diverse suite of distributions.

Example Suppose $X \sim \mathrm{Bin}(n, p)$ where p is modeled as $p \sim \mathrm{Beta}(\alpha, \beta)$. Then

$$
\begin{aligned}
P(X = k) &= \frac{1}{B(\alpha, \beta)} \int_0^1 \binom{n}{k} p^k (1-p)^{n-k} p^{\alpha-1} (1-p)^{\beta-1} \, \mathrm{d}p \\
&= \frac{1}{B(\alpha, \beta)} \binom{n}{k} \int_0^1 p^{k+\alpha-1} (1-p)^{n-k+\beta-1} \, \mathrm{d}p \\
&= \binom{n}{k} \frac{B(k+\alpha, n-k+\beta)}{B(\alpha, \beta)} \\
&= \binom{n}{k} \frac{\Gamma(k+\alpha)\Gamma(n-k+\beta)}{\Gamma(n+\alpha+\beta)} \times \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \\
&= \frac{\Gamma(\alpha+\beta)\Gamma(n+1)\Gamma(k+\alpha)\Gamma(n-k+\beta)}{\Gamma(k+1)\Gamma(n-k+1)\Gamma(\alpha)\Gamma(\beta)\Gamma(n+\alpha+\beta)}.
\end{aligned}
$$

46

$$P(X = k) = \frac{\Gamma(\alpha+\beta)\Gamma(n+1)\Gamma(k+\alpha)\Gamma(n-k+\beta)}{\Gamma(k+1)\Gamma(n-k+1)\Gamma(\alpha)\Gamma(\beta)\Gamma(n+\alpha+\beta)}.$$

When $\alpha = \beta = 1$, the above expression simplifies to

$$P(X = k) = \frac{1}{n+1}$$

which is the pmf of a discrete uniform rv on $\{0, 1, 2, \ldots, n\}$.

When $\alpha \to \infty$, $\beta \to \infty$ and $\alpha/(\alpha+\beta) \to p_* \in (0, 1)$ then

$$P(X = k) \to \binom{n}{k} p_*^k (1 - p_*)^{n-k}$$

which is the pdf for $\mathrm{Bin}(n, p_*)$.

47

Mean and variance of a beta rv

Proposition
Let $X \sim \mathrm{Beta}(\alpha, \beta)$. *Then*

$$\mathbb{E}[X] = \frac{\alpha}{\alpha + \beta}, \quad \mathrm{Var}[X] = \frac{\alpha\beta}{(\alpha + \beta)^2(\alpha + \beta + 1)}.$$

48

Proof Direct calculations and the definition of $B(\alpha, \beta)$ yield

$$
\begin{aligned}
\mathbb{E}[X] &= \frac{1}{B(\alpha,\beta)} \int_0^1 x^{\alpha}(1-x)^{\beta-1}\,\mathrm{d}x \\
&= \frac{B(\alpha+1,\beta)}{B(\alpha,\beta)} = \frac{\Gamma(\alpha+1)\Gamma(\beta)}{\Gamma(\alpha+\beta+1)} \times \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} = \frac{\alpha}{\alpha+\beta}.
\end{aligned}
$$

Similarly, we have

$$
\begin{aligned}
\mathbb{E}[X^2] &= \frac{1}{B(\alpha,\beta)} \int_0^1 x^{\alpha+1}(1-x)^{\beta-1}\,\mathrm{d}x \\
&= \frac{B(\alpha+2,\beta)}{B(\alpha,\beta)} = \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)}.
\end{aligned}
$$

More algebraic manipulations yield the expression for $\mathrm{Var}[X]$.

49

Higher order moments, as well as inverse and fractional moments of a Beta rv are also easy to derive (cf the gamma rv).

Proposition
Let $X \sim \mathrm{Beta}(\alpha, \beta)$. Then for all $c > -\alpha$

$$\mathbb{E}[X^c] = \frac{1}{B(\alpha,\beta)} \int_0^1 x^{c+\alpha-1}(1-x)^{\beta-1}\,\mathrm{d}x$$
$$= \frac{B(\alpha+c,\beta)}{B(\alpha,\beta)} = \frac{\Gamma(\alpha+c)\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\alpha+\beta+c)}.$$

50

85. Let X have a standard beta density with parameters α and β.

a. Verify the formula for $E(X)$ given in the section.

b. Compute $E[(1 - X)^m]$. If X represents the proportion of a substance consisting of a particular ingredient, what is the expected proportion that does not consist of this ingredient?

51

85. Let X have a standard beta density with parameters α and β.

a. Verify the formula for $E(X)$ given in the section.

b. Compute $E[(1-X)^m]$. If X represents the proportion of a substance consisting of a particular ingredient, what is the expected proportion that does not consist of this ingredient?

$1-X \sim \mathrm{Beta}(\beta,\alpha)$ and hence, by the previous proposition,

$$\mathbb{E}[(1-X)^m] = \frac{\Gamma(\beta+m)\Gamma(\alpha+\beta)}{\Gamma(\beta)\Gamma(\alpha+\beta+m)}.$$

51