Commonly Used Continuous Distributions **Notes 04**

Associated Reading: Wackerly 7, Chapter 4, Sections 4-8

In these notes, we shift from talking about discrete distributions to continuous distributions, highlighting the uniform, normal, gamma, and beta families. *Roughly* speaking, the difference between using discrete and continuous distributions is that theoretical considerations motivate the use of the former more than they do the use of the latter. For instance, tossing a coin is a Bernoulli process, and thus the number of successes in n trials *is* a binomial random variable. On the other hand, we can probably model the distribution of student heights at UPenn well using the normal distribution, but there is no theory that says that the true underlying distribution has to be normal.

This leads naturally to the question: how do we choose between families of continuous distributions when modelling phenomena? The fullest answer is that it is not necessarily easy, and to make any choice we have to know something about parameter estimation and hypothesis testing, topics for 36-226. So at this stage in the course, we are pretty much limited to simply introducing various distributions and their properties. Which we'll do now.

$\propto$

$Y \in [a, b]$.

The simplest continuous distribution is the *uniform distribution*, in which all values between $y = a$ and $y = b$ have equal probability density. In practice, the uniform distribution is used as much or more for simulating data (or generating random numbers) than for actual modeling.

THE UNIFORM DISTRIBUTION

Unif (a, b)

NOTATION: $Y \sim \text{Uniform}(a, b)$ a: LOWER BOUND, b: UPPER BOUND

PDF: $f(y) = \begin{cases} \frac{1}{b-a} & a \le y \le b \\ 0 & \text{OTHERWISE} \end{cases}$ $\Rightarrow \int_a^b f(y)\, dy = 1$

EXPECTED VALUE: $E[Y] = \mu =$ SEE BELOW

VARIANCE: $V[Y] = \sigma^2 =$ SEE BELOW

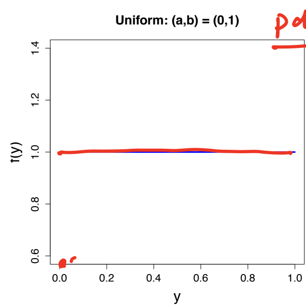
R FUNCTIONS: `dunif(y, a, b)` (PDF)

`punif(y, a, b)` (CDF)

`qunif(p, a, b)` (INVERSE CDF) quantile.

`runif(k, a, b)` (SIM. OF k UNIFORM R.V.'S)

Unif (0,1) pdf



f(y) general density.

① $\int f(y) dy = 1$ Domain

② $f(y) \ge 0$.

As might be expected, $E[Y]$ and $V[Y]$ are exceptionally easy to compute:

$$E[Y] = \int_a^b y f(y)\, dy = \int_a^b y \cdot \frac{1}{b-a}\, dy = \frac{b+a}{2}.$$

$$V[Y] = E[Y^2] - (E[Y])^2 = \int_a^b y^2 f(y)\, dy - \left(\frac{b+a}{2}\right)^2 = \frac{1}{12}(b-a)^2.$$

check this offline.

F(y) general cdf

① $\lim_{y \to \infty} F(y) = 1$

② $\lim_{y \to -\infty} F(y) = 0$

③ for $y_1 \le y_2$, $F(y_1) \le F(y_2)$.

$\boxed{\text{quantile:}}$ for $0 \le p \le 1$, the p-th quantile of r.v. Y. denoted by ϕ_p, is the smallest value st:

$F(\phi_p) = P(Y \le \phi_p) = p \Rightarrow \phi_p = F^{-1}(p).$

→ **EXAMPLE.** Wackerly 7, Exercise 4.47

4.47 The failure of a circuit board interrupts work that utilizes a computing system until a new board is delivered. The delivery time, Y, is uniformly distributed on the interval one to five days. The cost of a board failure and interruption includes the fixed cost c_0 of a new board and a cost that increases proportionally to Y^2. If C is the cost incurred, $C = c_0 + c_1 Y^2$.

a Find the probability that the delivery time exceeds two days.

b In terms of c_0 and c_1, find the expected cost associated with a single failed circuit board.

a). $P(Y \ge 2) = \int_2^5 f(y)\,dy = \int_2^5 \frac{1}{5-1}\,dy = \frac{3}{4}.$

$Y \sim \text{Unif}(1, 5).$

b). $E[C]$.

$= E[c_0 + c_1 Y^2]$

$= c_0 + c_1 E[Y^2]$

$= c_0 + c_1 \left(V[Y] + (E[Y])^2 \right)$

$= c_0 + c_1 \left[\frac{(5-1)^2}{12} + \left(\frac{5+1}{2}\right)^2 \right]$

$= c_0 + \frac{31}{3} c_1$

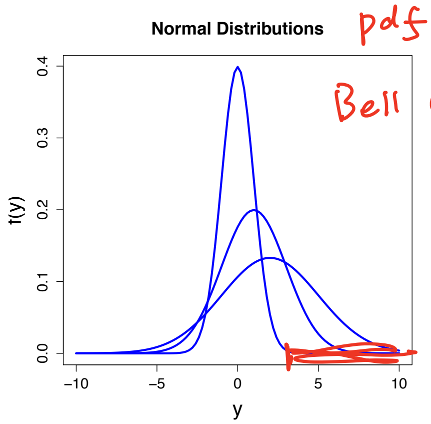
∼

Modeling Physical Phenomena

domain.
- $(-\infty, +\infty)$ → normal dist (Gaussian dist).
- $[0, +\infty)$ → Gamma dist
- $[0, 1]$. → Beta dist.

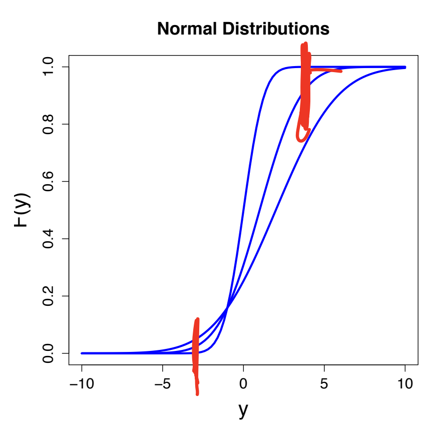
∼

The *normal distribution* is perhaps the most well-known family of probability distributions, for three reasons: (1) the observed data of many physical phenomena are at least approximately normal, (2) it is the limiting distribution of many other distributions and (3) it figures prominently in the Central Limit Theorem (Chapter 7).



pdf.

Bell curve. Shape.



cdf.

$E[Y^2] - (EY)^2$

$\int_{-\infty}^{\infty} y f(y)\,dy$

① $\int_{-\infty}^{+\infty} f(y)\,dy = 1$ ② $E[Y] = \mu$. ③ $V[Y] = \sigma^2$

$\int_{-\infty}^{+\infty} \exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right) dy = \sqrt{2\pi\sigma^2}$ $\quad \int_{-\infty} y \cdot \frac{1}{\sqrt{2\pi\sigma^2}} e^{\cdots} dy = \mu$ $\quad \int_{-\infty} y \cdot f(y)\, dy - \mu = \sigma$

$\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-\mu)^2}{2\sigma^2}}$

$N(\mu, \sigma^2)$.

The Normal Distribution

Notation: $Y \sim \mathrm{N}(\mu, \sigma^2)$ $\quad \mu$: mean, σ^2: variance (and not standard deviation!)

PDF: $f(y) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right)$ $\quad y \in (-\infty, \infty), \mu \in \mathbb{R}, \sigma \in \mathbb{R}^+$

$\exp(x) = e^x$

$\int_{-\infty}^{+\infty} y f(y)\, dy = \mu$

CDF: $F(y) = \frac{1}{2}\left[1 + \text{ERF}\left(\frac{y-\mu}{\sqrt{2}\sigma}\right)\right]$

where ERF(·) is the error function $\quad \text{ERF}(z) := \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt$

Expected Value: $E[Y] = \mu$ (by definition) Variance: $V[Y] = \sigma^2$ (by definition)

R Functions:
```
dnorm(y, mu, sigma) (PDF)
pnorm(y, mu, sigma) (CDF)
qnorm(p, mu, sigma) (Inverse CDF)
rnorm(k, mu, sigma) (Sim. of k Normal r.v.'s)
```

The CDF of the normal distribution includes the *error function*:[a]

$$\text{ERF}(z) := \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} dt.$$

$$F(y) = \int_{-\infty}^{y} f(x)\, dx = \int_{-\infty}^{y} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) dx.$$

let us $t = \frac{x-\mu}{\sqrt{2\sigma^2}}$

$$= \int_{-\infty}^{\frac{y-\mu}{\sqrt{2\sigma^2}}} \frac{1}{\sqrt{2\pi\sigma^2}} \exp(-t^2) \sqrt{2\sigma^2}\, dt$$

$$= \int_{-\infty}^{\frac{y-\mu}{\sqrt{2\sigma^2}}} \frac{1}{\sqrt{\pi}} \exp(-t^2)\, dt$$

$$= \int_{-\infty}^{0} \frac{1}{\sqrt{\pi}} \exp(-t^2)\, dt + \int_0^{\frac{y-\mu}{\sqrt{2\sigma^2}}} \frac{1}{\sqrt{\pi}} \exp(-t^2)\, dt$$

$$= \int_0^{\infty} \frac{1}{\sqrt{\pi}} \exp(-t^2)\, dt + \frac{1}{2} \text{ERF}\left(\frac{y-\mu}{\sqrt{2\sigma^2}}\right)$$

$\frac{1}{2} \text{ERF}(\infty)$ we know $= 1$.

$N(0,1)$.

Note that any normal random variable Y can be transformed to a *standard normal* random variable Z:

$Y \sim N(\mu, \sigma^2)$.

$V\left[\frac{Y-\mu}{\sigma}\right] = \frac{1}{\sigma^2} V[Y-\mu] = \frac{1}{\sigma^2} V[Y] = 1$

$V[cY] = c^2 V[Y]$.

$Z \triangleq \frac{Y-\mu}{\sigma} \sim N(0,1)$.

This is important for (historical) computational reasons:

$P(a \le Y \le b)$ $\quad N(0,1)$.

$= P\left(\frac{a-\mu}{\sigma} \le \frac{Y-\mu}{\sigma} \le \frac{b-\mu}{\sigma}\right)$.

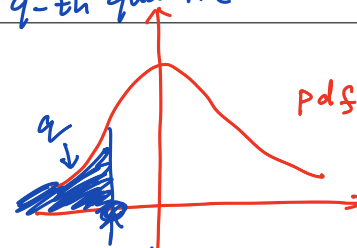
let us define $\Phi(z) = P(Z \le z)$.
cdf for $N(0,1)$

[a] You do not need to reproduce this derivation on a test.

$= \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$.

The Standard Normal Distribution

q-th quantile



c (is the q-th quantile).

$\boxed{\int_{-\infty}^{c} f(y)\,dy} = q.$

P(Y ≤ c)

Φ(c).

NOTATION: $Z \sim N(0,1)$ $\quad \mu = 0, \sigma^2 = 1$

PDF: $\boxed{f(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right)}$ $\quad z \in (-\infty, \infty)$

quantile.

CDF: $F(y) = \frac{1}{2}\left[1 + \cancel{\text{ERF}\left(\frac{z}{\sqrt{2}}\right)}\right] = \Phi(z)$ $\quad$ INVERSE CDF: $\Phi^{-1}(q) = z$

WHERE ERF(·) IS THE ERROR FUNCTION AND $q \in [0,1]$ IS A QUANTILE

EXPECTED VALUE: $E[Z] = 0$ $\quad$ VARIANCE: $V[Z] = 1$

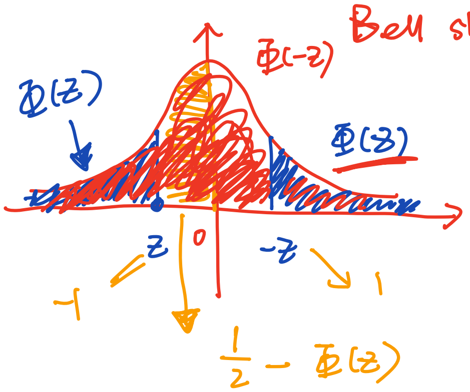
R FUNCTIONS: dnorm(z) (PDF) $\quad \mu = 0 \quad \sigma = 1$

pnorm(z) (CDF) ← $P(Z \le z)$

qnorm(p) (INVERSE CDF) ← $\Phi^{-1}(p)$

rnorm(k) (SIM. OF k STANDARD NORMAL R.V.'S)

Bell shape curve



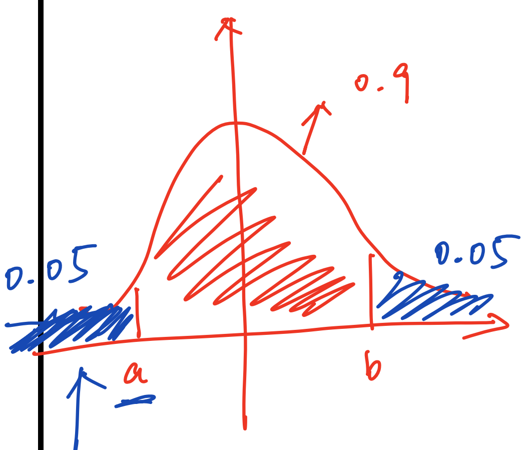
z<0, z>0.

$\boxed{\Phi(z) + \Phi(-z) = 1}$

by symmetry

notable numbers. $\quad$ N(0,1). $\quad$ pnorm(b) − pnorm(a)

(a, b)	$P(a \le Z \le b) = \Phi(b) - \Phi(a)$.
(−1, 1)	$\Phi(1) - \Phi(-1) = 0.6827$.
(−2, 2)	$\Phi(2) - \Phi(-2) = 0.9544$
(−3, 3)	0.9973.



$\Phi(a) = 0.05$.

$\Phi(b) - \Phi(a)$	(a, b) $\quad a = -b$.
0.9	$a = \text{qnorm}(0.05) = \Phi^{-1}(0.05)$. (−1.645, 1.645)
0.95	$a = \Phi^{-1}(0.025)$ $\boxed{(-1.960, 1.960)}$
0.99	(−2.576, 2.576).

→ **EXAMPLE.** Wackerly 7, Exercise 4.73

4.73 The width of bolts of fabric is normally distributed with mean 950 mm (millimeters) and standard deviation 10 mm.

a What is the probability that a randomly chosen bolt has a width of between 947 and 958 mm?

b What is the appropriate value for C such that a randomly chosen bolt has a width less than C with probability .8531?

Y: width of bolts of fabric.

$Y \sim N(950, 10^2)$ → variance.

a) $P(947 \le Y \le 958)$

$= P\left(\frac{947-950}{10} \le \boxed{\frac{Y-950}{10}} \le \frac{958-950}{10}\right)$ $\sim N(0,1)$.

cdf. $= \boxed{\Phi(0.8) - \Phi(-0.3)}$. exam answer.

$=$ pnorm(0.8) − pnorm(−0.3) $= \boxed{0.406}$. HW answer.

b) $P(Y \le c) = 0.8531$. find c.

$P\left(\boxed{\frac{Y-950}{10}} \le \frac{c-950}{10}\right)$. $\sim N(0,1)$

$\Phi\left(\frac{c-950}{10}\right) = 0.8531$

$\Rightarrow c = \boxed{950 + 10 \cdot \Phi^{-1}(0.8531)}$. exam.

$= 950 + 10 \cdot$ qnorm(0.8531)

$= \boxed{960.5}$ HW.

→ **EXAMPLE.** Wackerly 7, Exercise 4.75

4.75 A soft-drink machine can be regulated so that it discharges an average of μ ounces per cup. If the ounces of fill are normally distributed with standard deviation 0.3 ounce, give the setting for μ so that 8-ounce cups will overflow only 1% of the time.

Y: the ounces of fill.

$Y \sim N(\mu, 0.3^2)$

$\boxed{P(Y \ge 8)} = 0.01 \Rightarrow$ Goal: find μ.

$P\left(\boxed{\frac{Y-\mu}{0.3}} \ge \frac{8-\mu}{0.3}\right)$

$1 - \Phi\left(\frac{8-\mu}{0.3}\right) = 0.01$

$\Phi\left(\frac{8-\mu}{0.3}\right) = 0.99$.

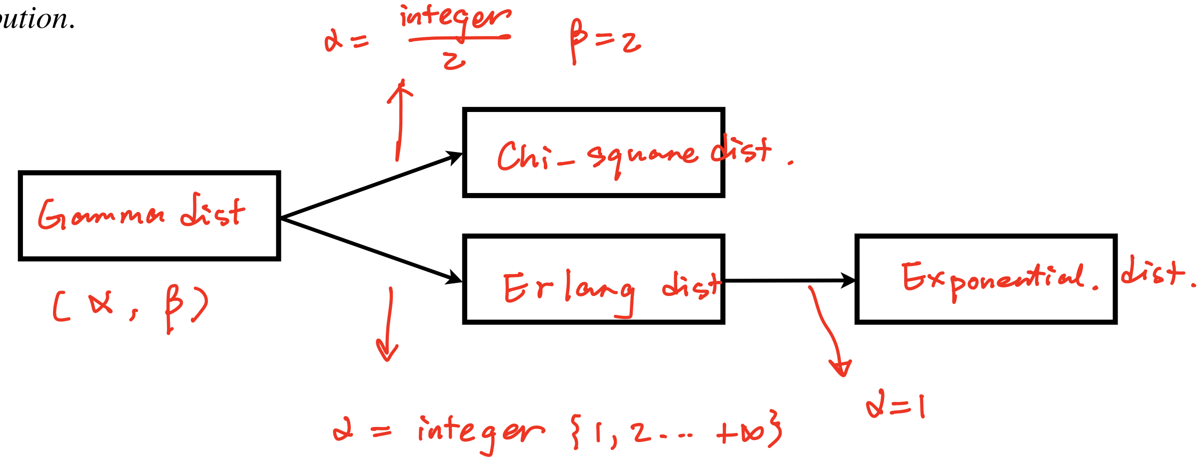
$\Rightarrow \mu = \boxed{8 - 0.3\,\Phi^{-1}(0.99)}$ exam

$= 8 - 0.3$ qnorm(0.99)

$= \boxed{7.302}$ HW.

$\propto$

The *gamma distribution* is a skew distribution used to model phenomena yielding non-negative random variables. For instance, we can model the time between events in a Poisson process (which cannot be less than zero) with a particular subclass of the gamma distribution called the *exponential distribution*.



$f(y) \propto y^{\alpha-1} \cdot e^{-\frac{y}{\beta}}$ scales / equal up to a constant.

$\int_0^\infty y^{\alpha-1} e^{-\frac{y}{\beta}} dy = \beta^\alpha \cdot \Gamma(\alpha)$

The Gamma Distribution

Notation: $Y \sim \text{Gamma}(\alpha, \beta)$ or $Y \sim \Gamma(\alpha, \beta)$ (α: shape parameter, β: scale parameter)

PDF: $f(y) = \begin{cases} \frac{y^{\alpha-1} e^{-y/\beta}}{\beta^\alpha \Gamma(\alpha)} & y \in [0, \infty), (\alpha, \beta) \in \mathbb{R}^+ \\ 0 & \text{otherwise} \end{cases}$ CDF: $F(y) = \frac{1}{\Gamma(\alpha)} \gamma(\alpha, \frac{y}{\beta})$

$\gamma(a, b) = \int_0^b t^{a-1} e^{-t} dt$

$b = \frac{y}{\beta}$

where $\gamma(\alpha, \frac{y}{\beta})$ is the lower incomplete gamma function and

Gamma Fuction $\Gamma(\alpha) = \int_0^\infty y^{\alpha-1} e^{-y} dy$

$\Gamma(n) = (n-1)!$ if $n = 1, 2, 3, \ldots$ prove this later.

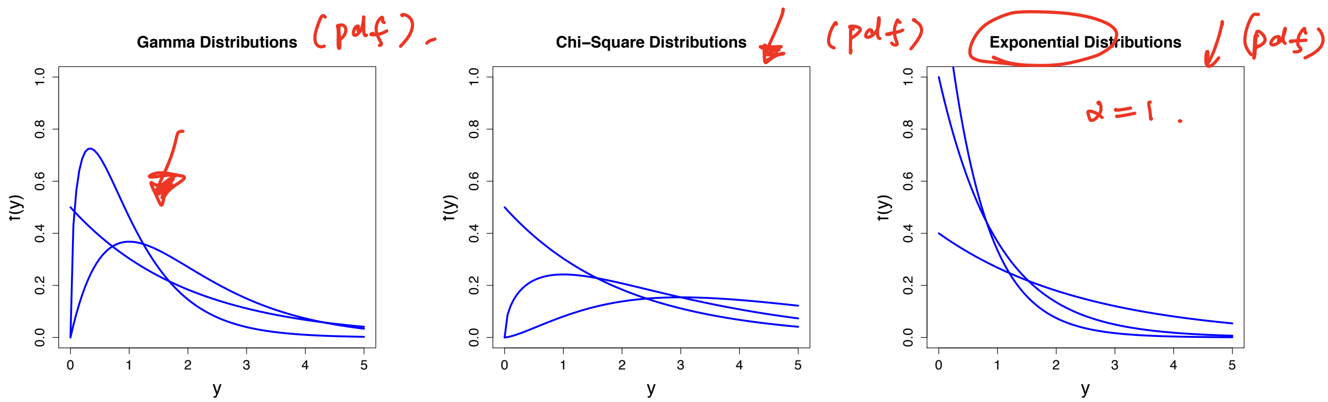
Expected Value: $E[Y] = \mu = \alpha\beta$ Variance: $V[Y] = \sigma^2 = \alpha\beta^2$ $\Rightarrow$ prove these two later.

R Functions:
- `dgamma(y, alpha, 1/beta)` (PDF)
- `pgamma(y, alpha, 1/beta)` (CDF)
- `qgamma(p, alpha, 1/beta)` (Inverse CDF)
- `rgamma(k, alpha, 1/beta)` (Sim. of k Gamma r.v.'s)

$Y \sim \text{Gamma}(\alpha, \beta)$

$f(y)$ pdf

$\text{dgamma}(y, \alpha, \frac{1}{\beta})$.



→ **EXAMPLE.** Wackerly 7, Exercise 4.81

$\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$

4.81 **a** If $\alpha > 0$, $\Gamma(\alpha)$ is defined by $\Gamma(\alpha) = \int_0^\infty y^{\alpha-1} e^{-y} dy$, show that $\Gamma(1) = 1$.

***b** If $\alpha > 1$, integrate by parts to prove that $\Gamma(\alpha) = (\alpha - 1)\Gamma(\alpha - 1)$.

a). $\Gamma(\alpha) = \int_0^\infty y^{\alpha-1} e^{-y} dy$.

$\alpha = 1$ $\Gamma(1) = \int_0^\infty y^0 \cdot e^{-y} dy$

$= \int_0^\infty e^{-y} dy$.

$= -e^{-y} \Big|_0^\infty = 1$

b). $\Gamma(\alpha) = \int_0^\infty y^{\alpha-1} e^{-y} dy$.

integration by parts:

$\int_a^b u\, dv = u \cdot v \Big|_a^b - \int_a^b v\, du$

$= \int_0^\infty y^{\alpha-1} d(-e^{-y})$

$= y^{\alpha-1} \cdot (-e^{-y}) \Big|_0^\infty - \int_0^\infty -e^{-y} d(y^{\alpha-1})$

$= 0 \int_0^\infty e^{-y} \cdot (\alpha-1) \cdot y^{\alpha-2} dy$.

$= (\alpha-1) \int_0^\infty e^{-y} \cdot y^{\alpha-2} dy$.

$\Gamma(\alpha-1)$

n is integer

$\Gamma(n) = (n-1)!$

$\Gamma(1) = 1$ ✓

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

$\Gamma(n) = (n-1)\Gamma(n-1)$

$= (n-1) \cdot (n-2)$

$\Gamma(n-2)$

$= (n-1)(n-2)\cdots\Gamma(1)$

$= (n-1)!$

Here we lay out the derivation of $E[Y]$ and $V[Y]$:

$$E[Y] = \int_0^\infty y\, \underbrace{f(y)}_{\text{pdf Gamma}(\alpha,\beta)}\, dy = \int_0^\infty y \cdot \left(\frac{y^{\alpha-1} \cdot e^{-\frac{y}{\beta}}}{\beta^\alpha \cdot \Gamma(\alpha)} \right) dy.$$

$$= \frac{1}{\beta^\alpha \Gamma(\alpha)} \cdot \boxed{\int_0^\infty y^\alpha e^{-\frac{y}{\beta}}\, dy}$$

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

Recall $\text{Gamma}(\alpha+1, \beta)$.

$$\underset{\text{pdf}}{f(y)} = \frac{y^{\alpha+1-1} e^{-\frac{y}{\beta}}}{\beta^{\alpha+1}\, \Gamma(\alpha+1)} = \frac{y^\alpha e^{-\frac{y}{\beta}}}{\beta^{\alpha+1} \cdot \Gamma(\alpha+1)}.$$

$$1 = \int_0^\infty f(y)\,dy \Rightarrow \int_0^\infty \frac{y^\alpha \cdot e^{-\frac{y}{\beta}}}{\beta^{\alpha+1} \cdot \Gamma(\alpha+1)}\, dy = 1$$

$$= \frac{\beta^{\alpha+1}\, \Gamma(\alpha+1)}{\beta^\alpha\, \Gamma(\alpha)}$$

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

(from part (b) above).

$$= \beta \cdot \alpha.$$

$\text{Gamma}(\alpha+2, \beta)$

$$f(y) = \frac{y^{\alpha+1} e^{-\frac{y}{\beta}}}{\beta^{\alpha+2}\, \Gamma(\alpha+2)}$$

$$E[Y^2] = \int_0^\infty y^2 \cdot \frac{y^{\alpha-1} \cdot e^{-\frac{y}{\beta}}}{\beta^\alpha \cdot \Gamma(\alpha)} \cdot dy.$$

- pull out the constants;
- match pdf for another Gamma dist.

$$= \frac{1}{\beta^\alpha \Gamma(\alpha)} \cdot \boxed{\int_0^\infty y^{\alpha+1} \cdot e^{-\frac{y}{\beta}}\, dy}$$

$f(y)$ integrates to 1 again.

$$= \frac{1}{\beta^\alpha \Gamma(\alpha)} \cdot \beta^{\alpha+2}\, \Gamma(\alpha+2)$$

$$\Gamma(\alpha+2) = (\alpha+1)\Gamma(\alpha+1) = (\alpha+1)\,\alpha\,\Gamma(\alpha)$$

$$= \alpha(\alpha+1) \cdot \beta^2$$

$$V[Y] = E[Y^2] - (EY)^2 = (\alpha^2+\alpha)\beta^2 - \alpha^2\beta^2 = \alpha\beta^2$$

→ **EXAMPLE.** Wackerly 7, Exercise 4.105

4.105 Four-week summer rainfall totals in a section of the Midwest United States have approximately a gamma distribution with $\alpha = 1.6$ and $\beta = 2.0$.

$Y \sim \text{Gamma}(\alpha = 1.6, \beta = 2.0)$.

a) $E[Y]$, $V[Y]$

b) $P[Y > 4]$.

a) $E[Y] = \alpha \cdot \beta = 1.6 \times 2 = 3.2.$

$V[Y] = \alpha\beta^2 = 6.4.$

b) $P(Y > 4)$

$= 1 - P(Y \le 4)$

$= 1 - F(4)$

$= 1 - \text{pgamma}(4, 1.6, \boxed{\tfrac{1}{2}})$ ($\uparrow \frac{1}{\beta}$)

Check this offline.

let us define lower incomplete gamma function.

$$\gamma(a, b) := \int_0^{\boxed{b}} y^{a-1} e^{-y}\, dy.$$

$$\Gamma(a) := \int_0^{\boxed{\infty}} y^{a-1} e^{-y}\, dy.$$

$$\underset{\text{cdf}}{F(y)} \overset{\text{evaluate}}{=} \frac{\gamma(\alpha, \frac{y}{\beta})}{\Gamma(\alpha)}$$

$Y \sim \text{Gamma}(\alpha, \beta)$.

$\boxed{= 0.290.}$ HW answer.

$P(Y>4) = 1 - F(4) = \boxed{1 - \frac{\gamma(1.6, \frac{4}{2})}{\Gamma(1.6)}}$ exam answer

The exponential distribution can be used to model the time difference between two successive events in a Poisson process. (The Erlang distribution generalizes this to the time between k events.)

$Y \sim \text{Expo} / \text{Exp}(\beta)$

$\boxed{E[Y] = \beta}$

$V[Y] = \beta^2.$

$\text{Exp}(\beta) = \text{Gamma}(\alpha = 1, \beta).$

pdf. $f(y) = \frac{y^{1-1} e^{-\frac{y}{\beta}}}{\beta^1 \cdot \underbrace{\Gamma(1)}_{=1}} = \frac{1}{\beta} e^{-\frac{y}{\beta}}$

R functions: pdf: dexp$(y, \frac{1}{\beta})$

cdf: pexp$(y, \frac{1}{\beta})$

qexp, rexp ...

→ **EXAMPLE.** Wackerly 7, Exercise 4.91

4.91 The operator of a pumping station has observed that demand for water during early afternoon hours has an approximately exponential distribution with mean 100 cfs (cubic feet per second).

$Y \sim \text{Exp}(\beta = 100)$

a Find the probability that the demand will exceed 200 cfs during the early afternoon on a randomly selected day.

b What water-pumping capacity should the station maintain during early afternoons so that the probability that demand will exceed capacity on a randomly selected day is only .01?

→ cdf of Exp(100)

a) $P(Y > 200) = 1 - \underbrace{P(Y \le 200)}_{\text{cdf}} = 1 - \text{pexp}(200, \frac{1}{100})$

$= \int_{200}^{\infty} \frac{1}{\beta} e^{-\frac{y}{\beta}} dy = -e^{-\frac{y}{\beta}} \Big|_{200}^{\infty} = e^{-2}$

b) $P(Y > c) = 0.01$

Goal is to find c.

$\int_c^{\infty} \frac{1}{\beta} e^{-\frac{y}{\beta}} dy.$

$= -e^{-\frac{y}{\beta}} \Big|_c^{\infty}$

$= e^{-\frac{c}{100}}$

$\Rightarrow e^{-\frac{c}{100}} = 0.01.$

$\Rightarrow c = 100 \ln(100)$

$= 460.5.$

The chi-square distribution is especially important in fitting models to observed data:

$Y \sim \chi^2(n) = \text{Gamma}(\alpha = \frac{n}{2}, \beta = 2).$

n is an integer.

R function: dchisq(y, n) pchisq(y, n)

qchisq. rchisq.

$E[Y] = n \quad V[Y] = 2n.$ $\underset{\text{pdf}}{f(y)} = \frac{y^{\frac{n}{2}-1} e^{-\frac{y}{2}}}{2^{\frac{n}{2}} \cdot \Gamma(\frac{n}{2})}$

A fun fact chi-square dist.

let $X_1\ X_2 \cdots X_n$ independently, identically distributed.

$X_i \sim N(\mu, \sigma^2).$ $\quad Z_i = \frac{X_i - \mu}{\sigma} \sim N(0,1).$

$\sum_{i=1}^{n} Z_i^2 \sim \chi^2(n)$ $\quad n$: # of degrees of freedom.

$E[\sum^n Z_i^2] \stackrel{?}{=} n \cdot E[Z_i^2] = n \cdot [V(Z_i) + (E Z_i)^2]$

$V(\sum_{i=1}^{n} z_i^2) = \sum_{i=1}^{n} V(z_i^2) \cdot \overset{?}{=} 2n$ $= n.$

check this offline.

The *beta distribution* is superficially related to the gamma distribution, with the most notable difference being that it is defined over the range $0 \le y \le 1$.

Does this range limit the use of the beta distribution?

If $X \in \boxed{[a, b]}$, then $Y = \frac{X-a}{b-a} \in [0, 1]$

compact interval

now we can model dist of Y using Beta dist.

$f(y) \propto \boxed{y^{\alpha-1}(1-y)^{\beta-1}}$ $y \in [0,1]$

THE BETA DISTRIBUTION

$B(y; \alpha, \beta) := \int_0^y z^{\alpha-1}(1-z)^{\beta-1}\,dz$

NOTATION: $Y \sim \text{Beta}(\alpha, \beta)$ α: SHAPE PARAMETER, β: SCALE PARAMETER

PDF: $f(y) = \begin{cases} \frac{y^{\alpha-1}(1-y)^{\beta-1}}{B(\alpha,\beta)} & y \in [0,1], (\alpha,\beta) \in \mathbb{R}^+ \\ 0 & \text{OTHERWISE} \end{cases}$ $\boxed{\text{CDF: } F(y) = \frac{B(y;\alpha,\beta)}{B(\alpha,\beta)}}$ $\leftarrow P(Y \le y)$

WHERE $B(y; \alpha, \beta)$ IS THE INCOMPLETE BETA FUNCTION AND

$\rightarrow \int_0^1 f(y)\,dy = \frac{\int_0^1 y^{\alpha-1}(1-y)^{\beta-1}dy}{B(\alpha,\beta)} = 1$

when α, β are both integers $=$

$B(\alpha, \beta) = \int_0^1 y^{\alpha-1}(1-y)^{\beta-1}dy = \Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha+\beta)$ if α integer $\Gamma(\alpha) = (\alpha-1)!$

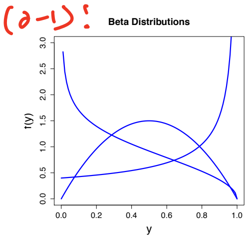
$\frac{(\alpha-1)!\,(\beta-1)!}{(\alpha+\beta-1)!}$

EXPECTED VALUE: $E[Y] = \mu = \alpha/(\alpha+\beta)$ (prove below)

VARIANCE: $V[Y] = \sigma^2 = \alpha\beta/[(\alpha+\beta)^2(\alpha+\beta+1)]$

R FUNCTIONS: `dbeta(y, alpha, beta)` (PDF)
`pbeta(y, alpha, beta)` (CDF)
`qbeta(y, alpha, beta)` (INVERSE CDF)
`rbeta(k, alpha, beta)` (SIM. OF k BETA R.V.'S)

⚠ in general ✗ close form for $B(\alpha, \beta)$



The CDF of the beta distribution is the incomplete beta function, which achieves a "simpler" form when $\alpha, \beta \in \mathbb{Z}^+$:

$F(y) = \int_0^y f(z)\,dz = \int_0^y \frac{z^{\alpha-1}(1-z)^{\beta-1}}{B(\alpha,\beta)}\,dz = \frac{1}{B(\alpha,\beta)} \int_0^y z^{\alpha-1}(1-z)^{\beta-1}\,dz$

$\int_0^y z^{\alpha-1}(1-z)^{\beta-1}dz$

$= \int_0^y (1-z)^{\beta-1}\, d\left(\frac{z^\alpha}{\alpha}\right)$

integration by part $= \frac{z^\alpha}{\alpha}(1-z)^{\beta-1}\Big|_0^y - \int_0^y \frac{z^\alpha}{\alpha}\, d(1-z)^{\beta-1}$

$= \frac{y^\alpha}{\alpha}(1-z)^{\alpha-1} - \int_0^y \frac{1}{\alpha} z^\alpha (\beta-1)(1-z)^{\beta-2} \cdot (-1)\,dz.$

$\vdots$ integrate by part.

$:= \boxed{\sum_{i=\alpha}^{n} \binom{n}{i} y^i (1-y)^{n-i}}$ $\alpha, \beta \in \mathbb{Z}^+$

where $n = \alpha+\beta-1$

check this offline.

if $Z \sim \text{Bin}(n, p = y)$

$= \boxed{P(Z \ge \alpha)}$

$= 1 - \underbrace{P(Z < \alpha)}_{\text{cdf for Bin.}}$

$\Rightarrow$ ① Beta dist & Binomial dist are fundamentally connected.

② in Bayesian statistics, Beta is conjugate prior to Binomial.

Here we lay out the derivation of $E[Y]$:

$Y \sim \text{Beta}(\alpha, \beta)$.

$$\boxed{E[Y]} = \int_0^1 y \cdot f(y)\, dy.$$

$$= \int_0^1 y \cdot \frac{y^{\alpha-1}(1-y)^{\beta-1}}{B(\alpha,\beta)}\, dy.$$

$$= \frac{1}{B(\alpha,\beta)} \cdot \boxed{\int_0^1 y^{\alpha}(1-y)^{\beta-1}\, dy.}$$

$B(\alpha+1, \beta)$

definition of Beta funtion.

$$= \frac{B(\alpha+1, \beta)}{B(\alpha,\beta)}$$

Beta dist $(\alpha+1, \beta)$.

$f(y) \propto y^{(\alpha+1-1)}(1-y)^{\beta}$

$$\frac{B(\alpha+1,\beta)}{B(\alpha,\beta)} = \frac{\dfrac{\Gamma(\alpha+1)\cancel{\Gamma(\beta)}}{\Gamma(\alpha+\beta+1)}}{\dfrac{\Gamma(\alpha)\cancel{\Gamma(\beta)}}{\Gamma(\alpha+\beta)}} = \boxed{\frac{\alpha}{\alpha+\beta}}$$

because

$\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$

exactly same procedure one can compute $E[Y^2]$.

$$V[Y] = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

Check this offline.

→ **EXAMPLE.** Wackerly 7, Exercise 4.131

4.131 Errors in measuring the time of arrival of a wave front from an acoustic source sometimes have an approximate beta distribution. Suppose that these errors, measured in microseconds, have approximately a beta distribution with $\alpha = 1$ and $\beta = 2$.

$Y \sim \text{Beta}(\alpha=1, \beta=2)$.

a What is the probability that the measurement error in a randomly selected instance is less than .5 μs?

b Give the mean and standard deviation of the measurement errors.

a) $$P(Y<0.5) = \int_0^{0.5} f(y)\,dy = \int_0^{0.5} \frac{y^{1-1}(1-y)^{2-1}}{B(1,2)}\,dy = \frac{1}{B(1,2)} \cdot \int_0^{0.5}(1-y)\,dy$$

$\alpha=1 \quad \beta=2$

$$= \frac{1}{\frac{\Gamma(1)\Gamma(2)}{\Gamma(3)}}\left(0.5 + \left.\frac{-y^2}{2}\right|_0^{0.5}\right)$$

$$= \frac{2!}{0!\ 1!}\left(0.5 - \frac{1}{8}\right) = \frac{3}{4}.$$

easier way

$$F(0.5) = P(Z \geq 1)$$ where $Z \sim \text{Bin}(n, p=y=0.5)$

$n = \alpha+\beta-1 = 2.$

$$= 1 - P(Z=0)$$

$$= 1 - \binom{2}{0} 0.5^0 (1-0.5)^2 = 0.75.$$

b). $$E[Y] = \frac{\alpha}{\alpha+\beta} = \frac{1}{1+2} = \frac{1}{3}.$$

$$V[Y] = \frac{\alpha\beta}{(\alpha+\beta)^2 \cdot (\alpha+\beta+1)} = \frac{1}{18} \qquad \sigma = \sqrt{V[Y]} = \frac{1}{3\sqrt{2}}.$$

→ **EXAMPLE.** Wackerly 7, Exercise 4.133

4.133 The proportion of time per day that all checkout counters in a supermarket are busy is a random variable Y with a density function given by

$$f(y) = \begin{cases} cy^2(1-y)^4, & 0 \le y \le 1, \\ 0, & \text{elsewhere.} \end{cases}$$

a Find the value of c that makes $f(y)$ a probability density function.

b Find $E(Y)$. (Use what you have learned about the beta-type distribution. Compare your answers to those obtained in Exercise 4.28.)

c Calculate the standard deviation of Y.

d **Applet Exercise** Use the applet *Beta Probabilities and Quantiles* to find $P(Y > \mu + 2\sigma)$.

$f(y) \propto y^{\alpha-1}(1-y)^{\beta-1}$

Beta$(\alpha=3, \beta=5)$

a) $c = \frac{1}{B(3,5)} = \frac{1}{\frac{\Gamma(3)\Gamma(5)}{\Gamma(8)}} = \frac{7!}{2!\cdot 4!} = \frac{7\cdot 6\cdot 5}{2} = 105.$

use $\Gamma(\alpha) = (\alpha-1)!$ for α: integer

b) $\mathbb{E}[Y] = \frac{\alpha}{\alpha+\beta} = \frac{3}{8}$

c) $V[Y] = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} = \frac{5}{192}. \Rightarrow \sigma = \sqrt{\frac{5}{192}}.$

d) $P(Y > \mu + 2\sigma)$

$= P\left(Y \geq \frac{3}{8} + 2\cdot\sqrt{\frac{5}{192}}\right).$

$= 1 - \text{pbeta}\left(\frac{3}{8} + 2\sqrt{\frac{5}{192}}, \alpha=3, \beta=5\right).$

$\boxed{= 0.030}$ HW answer.

Suppose. $B_1\ B_2 \cdots B_n$ form partition of S.

$\forall$ set A: $P(A) = \sum_{i=1}^{n} P(A \cap B_i)$ ... ② $\bigcup_{i=1}^{n} B_i = S$.

We concluded the last set of notes by demonstrating how the Law of Total Probability has a role in probabilistic modeling outside of when we work with sample spaces. In that demonstration, we assumed that we were working with two discrete distributions. Here, we generalize to the case where either one or both of the distributions is continuous.

There are three cases: X. Y. discrete / continuous.

1) The conditional distribution is discrete, and the unconditional distribution is continuous:

$P(Y|X)$ $f(x)$

- $P(Y=y) = \int_{D_x} P(Y|x) f(x)\,dx$

2) The conditional distribution is continuous, and the unconditional distribution is discrete:

$f(y|x)$ $P(x)$

- $f(y) = \sum_{x} f(y|x) P(x)$

$P(a \le Y \le b) = \int_a^b f(y)\,dy = \int_a^b \sum_x f(y|x)\, p(x)\, dy$

3) Both distributions are continuous:

- $f(y) = \int_{D_x} f(y|x) f(x)\,dx$

$P(a \le Y \le b) = \int_a^b \int_{D_x} f(y|x) f(x)\,dx\,dy$

→ **EXAMPLE.** Consider the following two-step experiment. First, X is drawn from a beta distribution with parameters α and β. Then Y is drawn from a binomial distribution with the number of trials being n and the probability of success being X. Determine $P[Y = y]$.

$X \sim \text{Beta}(\alpha, \beta)$.

$Y|X \sim \text{Bin}(n, X)$.

$$\boxed{P(Y=y)} = \int_{D_x} P(Y|X=x) f(x)\,dx$$

pmf

$$= \int_{D_x} \underbrace{\binom{n}{y} x^y (1-x)^{n-y}}_{\text{pmf for Bin}(n,x)} \cdot \underbrace{\frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha,\beta)}}_{\text{pdf Beta}(\alpha,\beta)} dx$$

$$= \frac{\binom{n}{y}}{B(\alpha,\beta)} \boxed{\int_0^1 x^{y+\alpha-1} (1-x)^{n+\beta-y-1}\,dx}.$$

$\Downarrow$ $B(y+\alpha, n+\beta-y)$.

$$\boxed{\binom{n}{y} B(y+\alpha, n+\beta-y)}$$

$$\boxed{B(\alpha, \beta).}$$

pmf.