Continuous Probability Distributions with R

Marianne Huebner, Michigan State University

2015-05-01

### Normal distribution

Average height of US males are approximately normal distributed with mean 69.1 inches and standard deviation 2.9 inches.

*Cumulative distribution function:* probability that a man is less than 70 inches tall.

pnorm(70, mean=69.1, sd=2.9)

## [1] 0.6218506

# What is the probability of a man being taller than 70 inches?

1-pnorm(70, mean=69.1, sd=2.9)

## [1] 0.3781494

# Calculate P(66< X <72) for X ~ Normal(69.9, 2.9)

pnorm(72, mean=69.1, sd=2.9)-pnorm(66, mean=69.1, sd=2.9)

## [1] 0.6988021

*Inverse cumulative distribution function.*

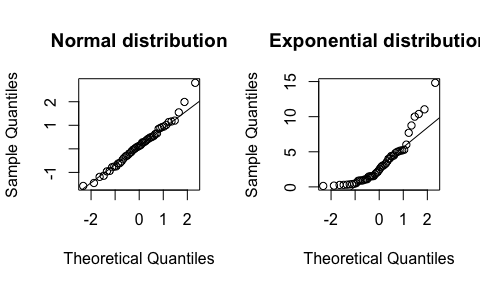
qnorm(0.9,mean=69.1, sd=2.9)

## [1] 72.8165

There is a 90% chance that a man is 72.8164995 inches tall or less.

Do the data come from a normal probability distribution?

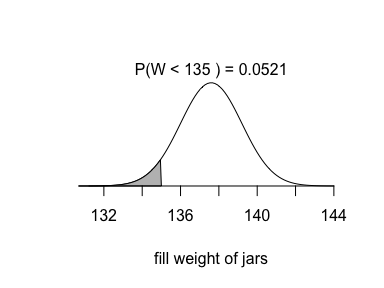
# generate normally distributed data  
z<-rnorm(50)  
# generate random numbers from a skewed distribution  
y<-rexp(50, rate=1/3)  
  
par(mfrow=c(1,2))  
qqnorm(z, main="Normal distribution")  
qqline(z)  
  
qqnorm(y, main="Exponential distribution")  
qqline(y)



par(mfrow=c(1,1))

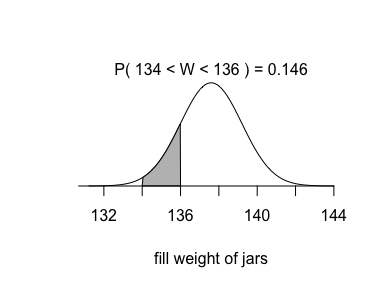
The fill weight of baby food jars follows a normal distribution with mean 137.6 grams and standard deviation of 1.6 grams. The label weight on the jars is 135.0 grams. *Reference: J. Fisher. Computer assisted net weight control. Quality Progress, June 1983.*

mean=137.6; sd=1.6  
lb=100; ub=135  
  
# draw a density curve for this normal distribution  
x <- seq(-4,4,length=100)\*sd + mean  
px <- dnorm(x,mean,sd)  
  
plot(x, px, type="n", xlab="fill weight of jars", ylab="", axes=FALSE)  
  
# shade the area for which the probability is calculated  
i <- x >= lb & x <= ub  
lines(x, px)  
polygon(c(lb,x[i],ub), c(0,px[i],0), col="gray")   
  
#Calculate the probability P(W<135) and paste it into the figure  
  
area <- pnorm(ub, mean, sd)  
result <- paste("P(W <",ub,") =", signif(area, digits=3))  
mtext(result,3)  
# add axes ticks and labels  
axis(1, at=seq(130, 145, 2), pos=0)



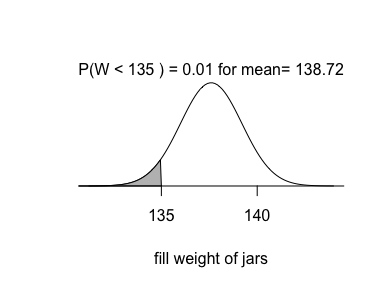
Create a figure and shade it for the probability

lb=134; ub=136  
  
plot(x, px, type="n", xlab="fill weight of jars", ylab="", axes=FALSE)  
  
# shade the area for which the probability is calculated  
i <- x >= lb & x <= ub  
lines(x, px)  
polygon(c(lb,x[i],ub), c(0,px[i],0), col="gray")   
  
  
area <- pnorm(ub, mean, sd) - pnorm(lb, mean, sd)  
result <- paste("P(",lb,"< W <",ub,") =", signif(area, digits=3))  
mtext(result,3)  
# add axes ticks and labels  
axis(1, at=seq(130, 145, 2), pos=0)



Unknown mean: create a figure and shade it for the probability

sd=1.6  
mean=135-qnorm(0.01)\*sd  
lb=100; ub=135  
  
plot(x, px, type="n", xlab="fill weight of jars", ylab="", axes=FALSE)  
  
# shade the area for which the probability is calculated  
i <- x >= lb & x <= ub  
lines(x, px)  
polygon(c(lb,x[i],ub), c(0,px[i],0), col="gray")   
  
  
area <- pnorm(ub, mean, sd)   
result <- paste("P(W <",ub,") =", signif(area, digits=3), "for mean=", round(mean,2))  
mtext(result,3)  
# add axes ticks and labels  
axis(1, at=seq(130, 145, 5), pos=0)



### Exponential distribution

Sixty-watt light bulbs have an average life of 1000 hours. The probability distribution of the lifetime of such a light bulb is exponential with rate=0.001

*Cumulative distribution function:* probability that the lifetime is less than 100 hours.

pexp(100, rate=1/1000)

## [1] 0.09516258

What is the probability that the light bulb lasts at least 900 hours?

1-pexp(900, 0.001)

## [1] 0.4065697

Calculate for

pexp(1200, rate=0.001)-pexp(900, rate=0.001)

## [1] 0.1053754

*Inverse cumulative distribution function.*

qexp(0.5,rate=0.001)

## [1] 693.1472

The median lifetime of such a bulb is 693.1471806 hours.

The top 20% of the light bulbs last at least how many hours?

qexp(0.8,rate=0.001)

## [1] 1609.438