Quantiles

Prof. Eric A. Suess

Quantiles

In base R there are 4 functions available for each probability model.

For example, *pnorm()*, *dnorm()*, *qnorm()* and there is a function to generate random values from the model *rnorm()*.

Probability Model

pnorm

To compute probabilities from the standard normal distribution. This function is equivalent to using a Normal Table in a book. It computes values of the CDF, cummulative probabilities.

$$F(z) = P(Z \leq z) = pnorm(z)$$

```
z <- 1.96

pnorm(z)  # Answer should be .975
```

```
[1] 0.9750021
```

$$P(-z \leq Z \leq z) = P(Z \leq z) - P(Z \leq -z) = F(z) - F(-z) =$$

$$pnorm(z) - pnorm(-z) = diff(pnorm(c(-z, z)))$$

```r
z <- c(-1.96, 1.96)

pnorm(z)
```

```
[1] 0.0249979 0.9750021
```

```r
diff(pnorm(z))  # computes the difference of the last value from the previous value.
```

```
[1] 0.9500042
```

dnorm

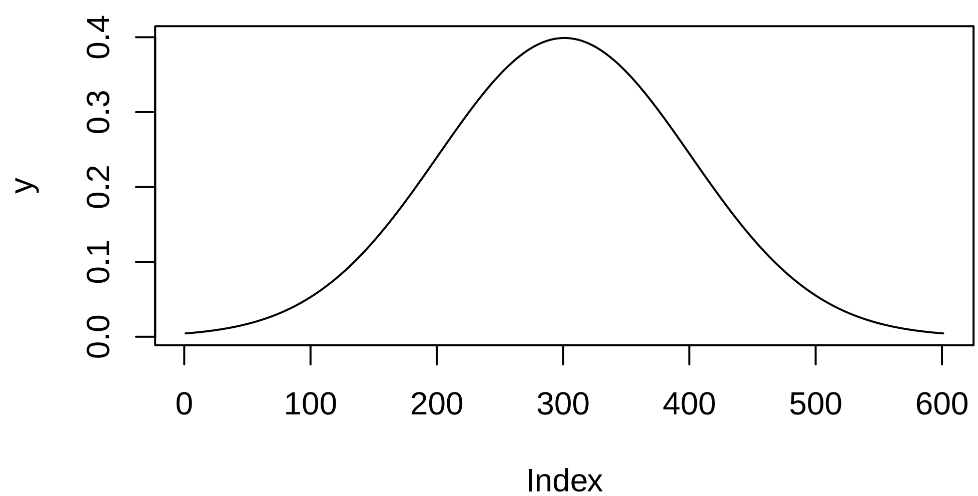
To compute the heights of the standard normal density, so we can plot the density curve.

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

```r
z <- seq(-3,3,.01)
head(z)
```

```
[1] -3.00 -2.99 -2.98 -2.97 -2.96 -2.95
```

```r
y <- dnorm(z)

plot(y, type = "l")
```



Note that we can integrate the *dnorm()* function to determine the area under the density function.

$$\int_{-z}^{z} f(z)dz = \int_{-z}^{z} \frac{1}{\sqrt{2\pi}} e^{-z^2} dz = integrate(dnorm, -z, z)$$

```r
integrate(dnorm, -1.96, 1.96)
```

```
0.9500042 with absolute error < 1e-11
```

qnorm

The quantile function for the normal is like using a T table in a book. Note that a quantile is the decimal equivalent to the percentile. So **percentile = 100 * quantile**.

$$F^{-1}(q) = qnorm(q)$$

Find the 95th quantile.

```r
q <- .975

qnorm(q)
```

```
[1] 1.959964
```

```r
round(qnorm(q), 2)
```

```
[1] 1.96
```

Notice that the *pnorm()* and *qnorm()* functions are inverses. The *pnorm()* function is the value of the CDF *F(z)* and *qnorm()* is the value of the inverse of the CDF $F^{-1}(q)$.

```r
z <- 1.96

pnorm(z)
```

```
[1] 0.9750021
```

```r
qnorm(pnorm(z))
```

```
[1] 1.96
```

```r
q <- .975

qnorm(q)
```

```
[1] 1.959964
```

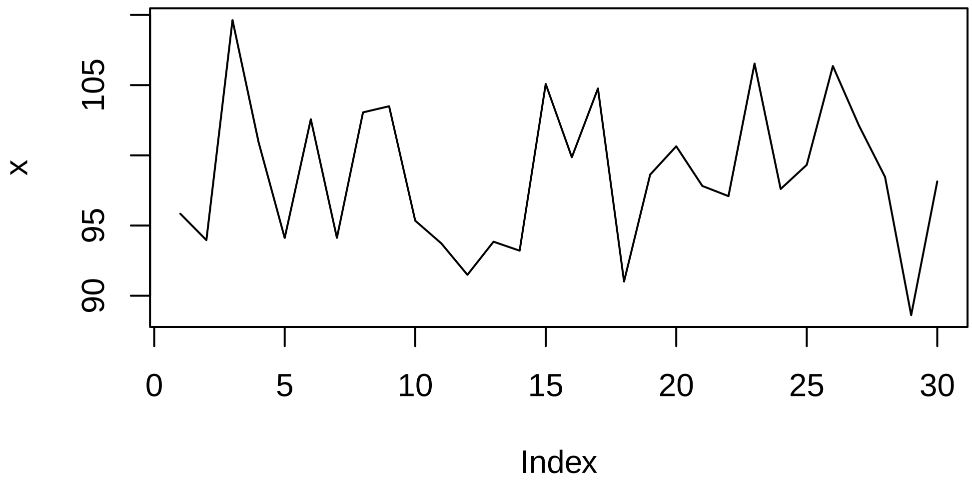
```r
pnorm(qnorm(q))
```

```
[1] 0.975
```

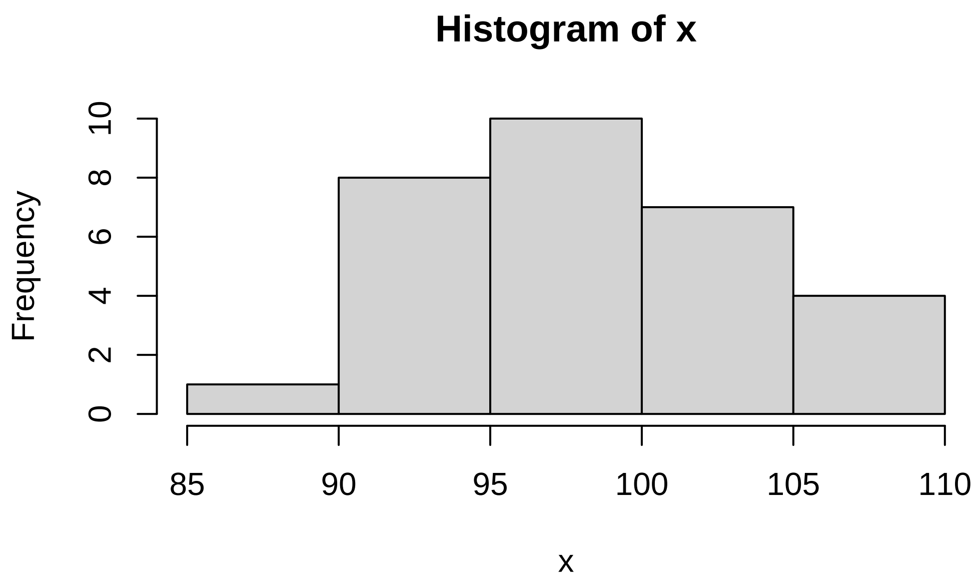
rnorm

To generate a **random sample** from the Normal distribution.

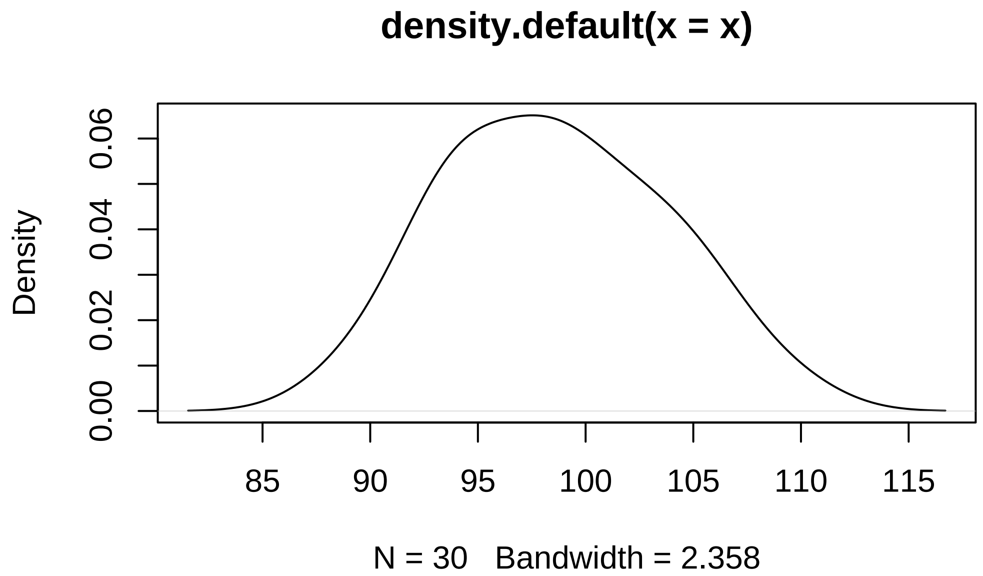
```r
n <- 30

mu <- 100
sigma <- 5

x <- rnorm(n, mu, sigma)

plot(x, type = "l")
```




```
hist(x)
```




```
plot(density(x))
```



Statistics can be computed from a random sample. Note that the *quantile()* function returns a value from a sample, just like the *mean()* function.

```
mean(x)
```

```
[1] 98.57989
```

```
sd(x)
```

```
[1] 5.172189
```

```
median(x)
```

```
[1] 98.3012
```

```
range(x)
```

```
[1]  88.61487 109.63272
```

```
IQR(x)
```

```
[1] 8.340737
```

```r
q <- .5

quantile(x)
```

```
      0%       25%       50%       75%      100% 
88.61487  94.11730  98.30120 102.45804 109.63272 
```

```r
min(x)
```

```
[1] 88.61487
```

```r
max(x)
```

```
[1] 109.6327
```