PROBABILITY DISTRIBUTIONS

PROF. DR. SALMAI QARI

OUTLINE

1. Distribution functions in R
2. Distribution functions for Normal Distribution: **pnorm, qnorm, dnorm** commands
    - definition and example: pnorm
    - definition and example: qnorm
    - definition and example: dnorm
3. Try the following:

```
x <- runif(2000)
y <- qnorm(x)
hist(y)
```

1. DISTRIBUTIONS IN R

Every distribution that R handles has four functions. There is a root name, for e.g. the root name for the normal distribution is **norm**. This root is prefixed by one of following letters:

- **p** for **probability**, the cumulative distribution function
- **q** for **quantile**, the inverse cumulative distribution function
- **d** for **density**, the density function or probability function
- **r** for **random**, a random variable having the specified distribution

2. DISTRIBUTION FUNCTIONS FOR NORMAL DISTRIBUTION

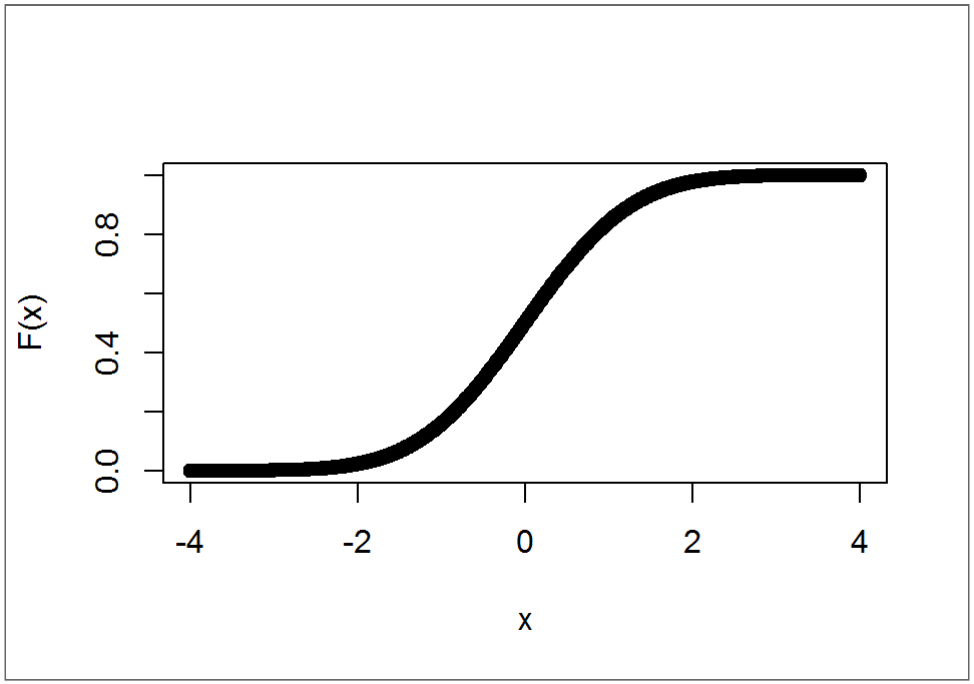
- dnorm(x, mean = 0, sd = 1, log = FALSE)
- pnorm(q, mean = 0, sd = 1, lower.tail = TRUE, log.p = FALSE)
- qnorm(p, mean = 0, sd = 1, lower.tail = TRUE, log.p = FALSE)
- rnorm(n, mean = 0, sd = 1)

PNORM DEFINITION

pnorm is the R function that calculates the **cumulative distribution function**

$F(x) = P(X <= x)$ where X follows a normal distribution.

```
xseq = seq(-4,4,0.01)
plot(x=xseq, y=pnorm(xseq), xlab="x", ylab="F(x)")
```



PNORM EXAMPLE

Question: Suppose the test scores of a college entrance exams are normally distributed with mean 72 and standard deviation 15.2 What is the probability that a randomly chosen candidate' score is higher than 84?

PNORM EXAMPLE: SOLUTION

$X \sim N(\mu = 72, \sigma = 15.2)$. What is $P(X > 84)$?

```
pnorm(84, mean=72, sd=15.2, lower.tail=FALSE)
```

```
## [1] 0.2149176
```

```
1 - pnorm(84, mean=72, sd=15.2)
```

```
## [1] 0.2149176
```

CALCULATING THE SAME BY SIMULATION (MONTE CARLO)

```
#sampling 100000 times
my.sample <- rnorm(100000, mean=72, sd=15.2)
#estimate probability
mean( my.sample > 84)
```

```
## [1] 0.2146
```

once again, exact solution:

```
1 - pnorm(84, mean=72, sd=15.2)
```

```
## [1] 0.2149176
```

QNORM DEFINITION

Inverse lookup qnorm is the R function that calculates the quantile function / inverse cumulative distribution function F^-1 of the normal distribution. The cumulative distribution function and the inverse cumulative distribution function are related by

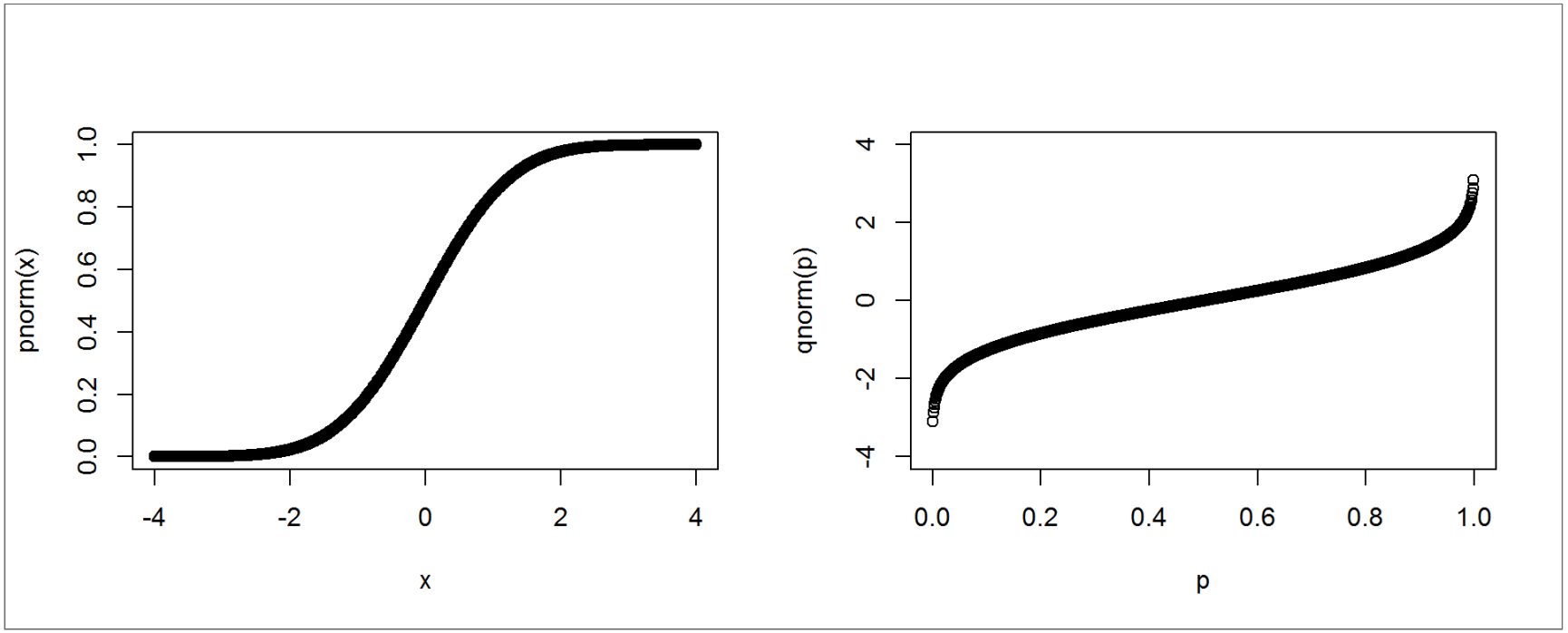
$$p = F(x)$$

$$x = F^{-1}(p)$$

So given a number p between 0 and 1, qnorm looks up the p-th quantile of the normal distribution.

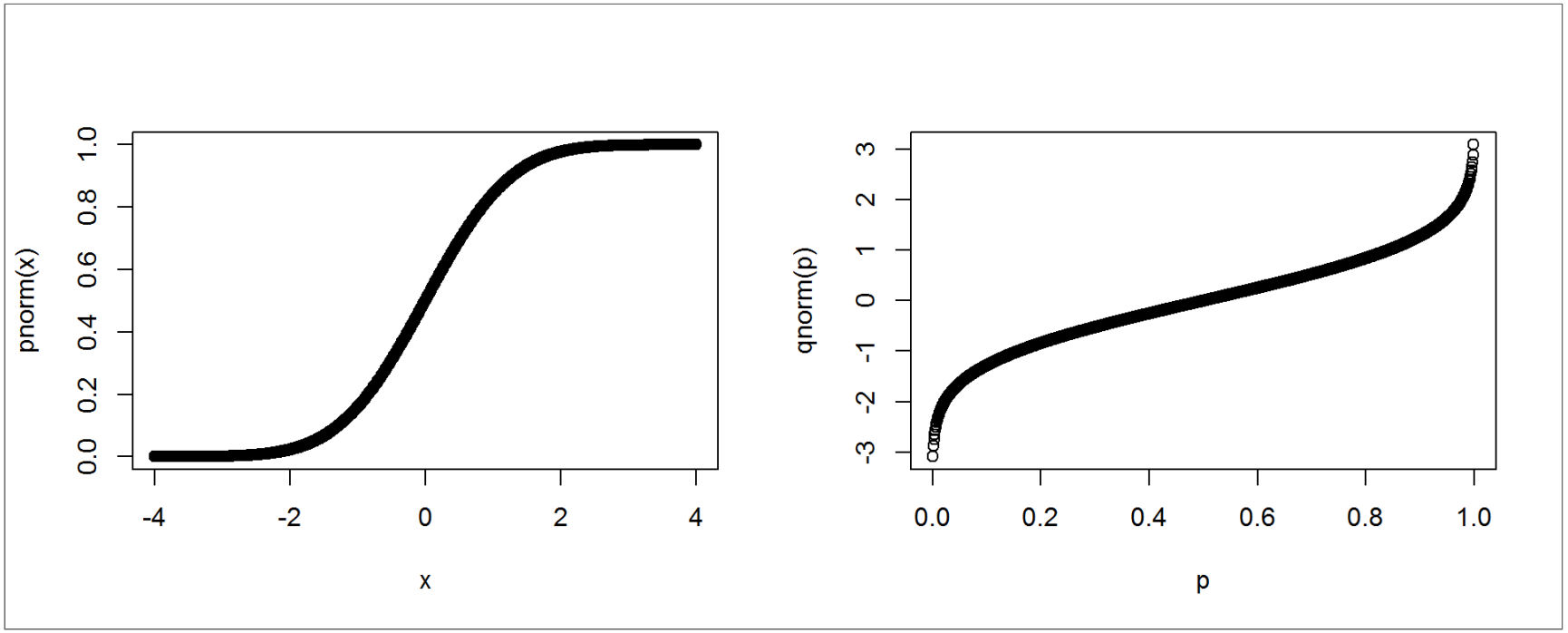
PROBABILITY FUNCTION (LHS); QUANTILE FUNCTION(RHS)

```
par(mfrow=c(1,2))
x = seq(-4,4,0.01); plot(x, pnorm(x));
p = seq(0,1,0.001); plot(y=qnorm(p), x=p, ylim=c(-4,4))
```



PROBABILITY FUNCTION (LHS); QUANTILE FUNCTION(RHS)

```
par(mfrow=c(1,2))
x = seq(-4,4,0.01); plot(x, pnorm(x))
p = seq(0,1,0.001); plot(p, qnorm(p))
```



QNORM EXAMPLE

Question: Suppose IQ scores are normally distributed with mean 100 and standard deviation 15. What is the 0.95-quantile of the distribution of IQ scores?

```
qnorm(0.95, mean=100, sd=15)
```

```
## [1] 124.6728
```

```
qnorm(0.05, mean=100, sd=15, lower.tail = F)
```

```
## [1] 124.6728
```

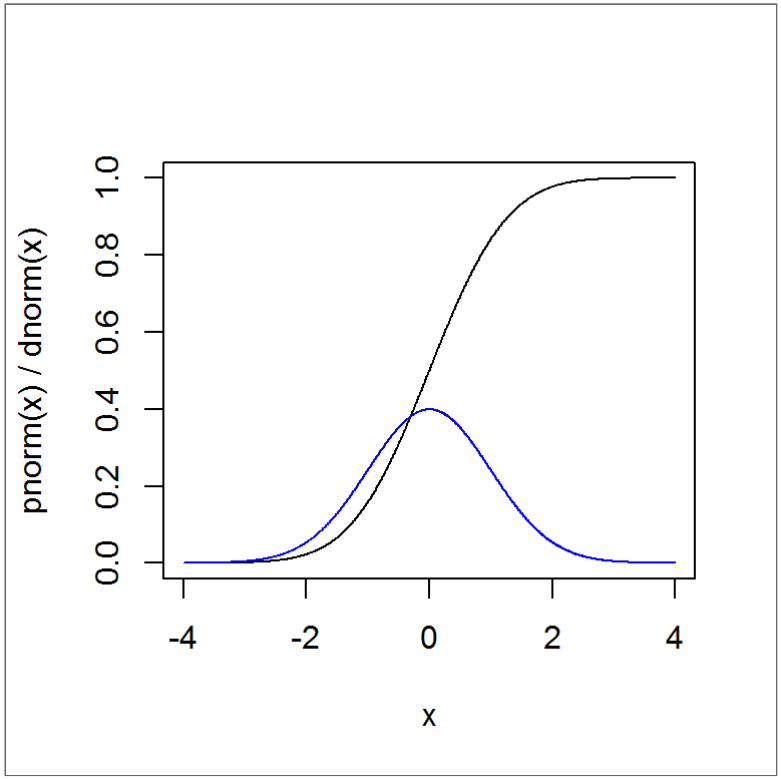
```
pnorm(124.6728, mean = 100, sd=15)
```

```
## [1] 0.95
```

DNORM DEFINITION

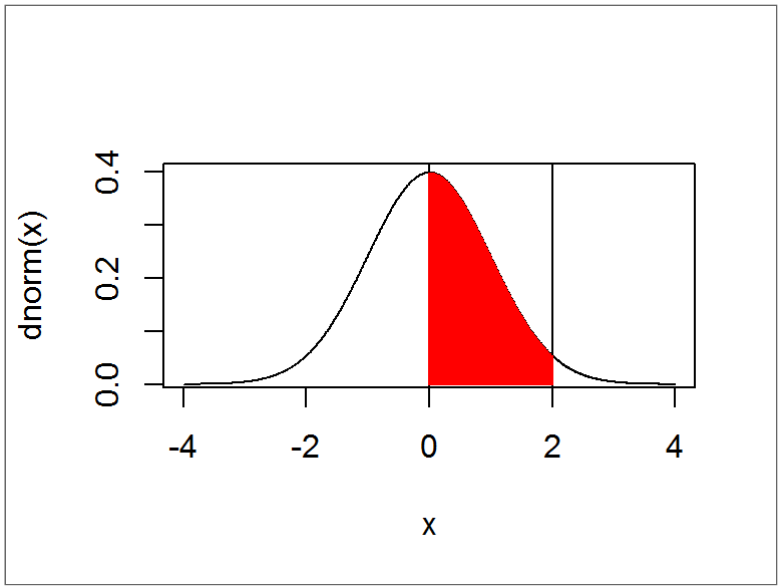
dnorm is the R function that calculates **probability density** function of the **normal distribution**. The **probability density** function of a continuous distribution is defined as the **derivative** of the **cumulative distribution** function.

```
x = seq(-4,4,0.01); plot(x, pnorm(x), type="l", ylab="pnorm(x) / dnorm(x)"); lines(x,dnorm(x),
         type="l", col="blue")
```


pnorm(x) / dnorm(x)
0.0
0.2
0.4
0.6
0.8
1.0
-4
-2
0
2
4
x

DNORM EXPLANATION

- Let $X \sim N(0,1)$. What is $P(0 \leq X \leq 2)$?

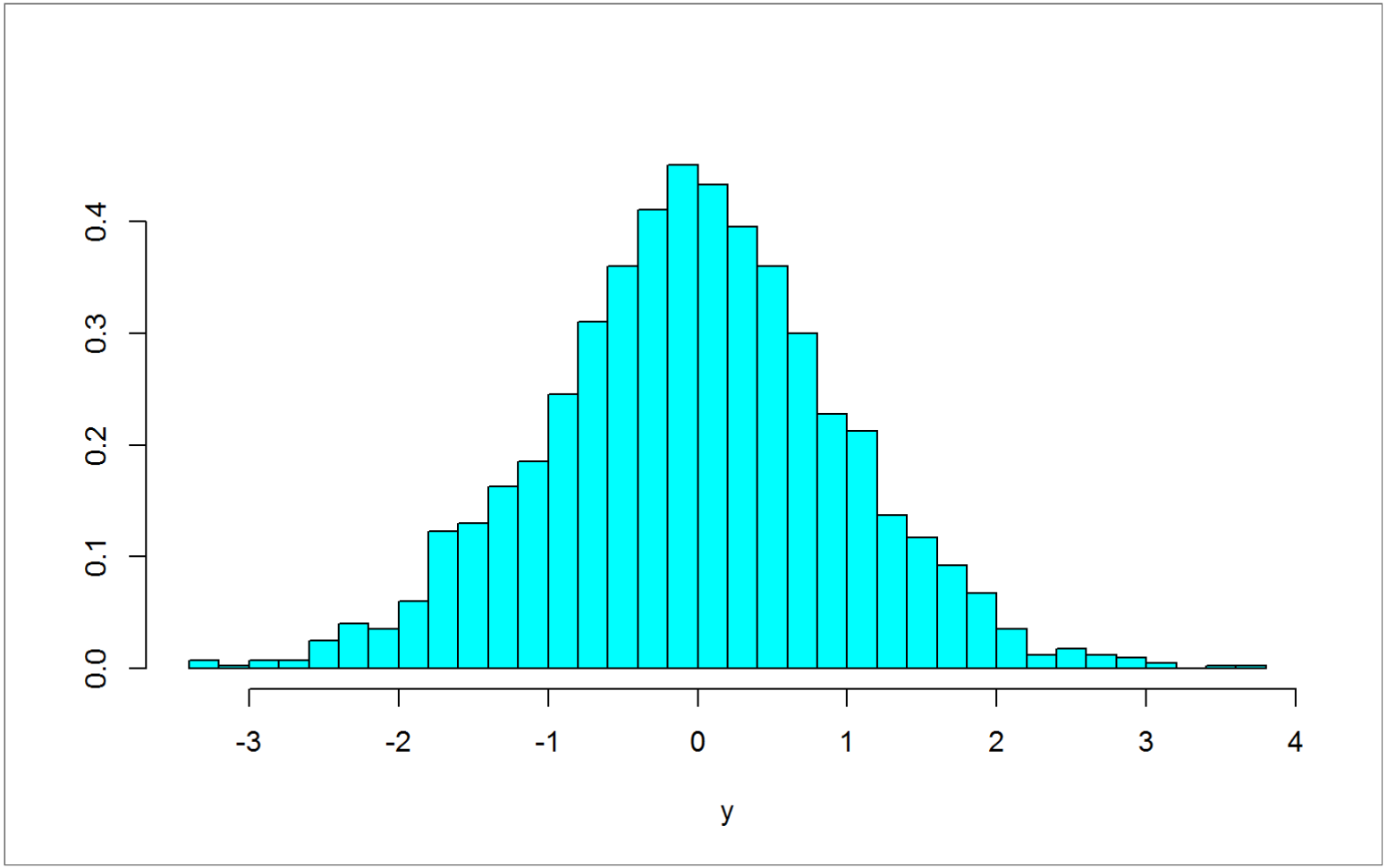


3. TRY THE FOLLOWING:

```
require(MASS)
```

```
## Loading required package: MASS
```

```
x <- runif(2000)
y <- qnorm(x)
truehist(y)
```


0.0
0.1
0.2
0.3
0.4
-3
-2
-1
0
1
2
3
4
y

SOURCES

- Coursera, Johns Hopkins University - R Programming, 2016
- Khan Academy,Random variables and probability distributions **https://www.khanacademy.org/math/probability/random-variables-topic/random-variables-prob-dist/v/probability-density-functions**
- **http://www.r-bloggers.com/normal-distribution-functions/**
- R documentation
- **http://www.stat.umn.edu/geyer/old/5101/rlook.html**
- **http://www.r-tutor.com/elementary-statistics/probability-distributions/normal-distribution**
- Wolfram, **http://mathworld.wolfram.com/ProbabilityDensityFunction.htm**