

# Overview document

## Contents

Libraries Used . . . . .	1
<b>Univariate descriptive Statistics</b>	<b>2</b>
Definition of moments. . . . .	2
Central Limit theorem. . . . .	3
Normal variable . . . . .	4
Law of large numbers . . . . .	6
Multivariate extension . . . . .	6
Calculating Sample Variance . . . . .	6
Combining variances. . . . .	6
Reading and estimating from graphs . . . . .	6
<b>Probability Distributions.</b>	<b>6</b>
pdf,cdf, icdf . . . . .	6
<b>Random Variables and Expectation</b>	<b>8</b>
<b>Null Hypothesis Sampling Testing.</b>	<b>8</b>

## Libraries Used

```
library(gmodels)
library(PSYC201)
library(ggplot2)
library(dplyr)
library(magrittr)
library(moments)
```

# Univariate descriptive Statistics

## Definition of moments.

Point of clarification: Moments refers just the the set of numbers that describe a set of points, it makes sense in a physical system or to probability. They are defined according to the following function.

$$\mu'_i = \int_a^b (x - c)^i f(x) dx$$

For probability  $f(x) = p(x)$  so :

$$\mu'_i = \int_a^b (x - c)^i p(x) dx$$

As  $i$  increases, you move up moments. The total number of moments as  $i$  goes to inf completely describes a set of points. *Seems like you need might only need  $n$  moments, but what do I know.* In general, each moment is referred to as  $\mu_i$ , my guess is that using  $\mu$  to represent mean is a convention to drop the 1.

Zeroeth Moment- the probability mass function

$$M_0 = \mu_0 = \Omega = 1$$

First moment, the mean

$$M_1 = \mu_1 = \mu_x = \int_a^b (x - c)^1 f(x) dx = \int_a^b x f(x) dx = E[X]$$

Second Moment, variance. Note that  $\sigma$  is the standard deviation.

$$M_2 = \mu_2 = \sigma_x^2 = \int_a^b (x - c)^2 f(x) dx = E[(X - E[X])^2] = E[X^2] - E[X]^2$$

Confirming with R

```
# need to adjust for sample calculation by subtracting 0, but not identitcal
x <- rnorm(200, 0, 4)
var(x)
```

```
## [1] 16.54169
```

```
sum((x^2))/(length(x) - 1) - mean(x)^2
```

```
## [1] 16.54298
```

This continues for the definition of moment, but using these in an non-normalized fashion doesn't appear normal

Third moment

$$M_3 = \mu_3 = \int_a^b (x - c)^3 f(x) dx$$

Forth Moment

$$M_4 = \mu_4 = \int_a^b (x - c)^4 f(x) dx$$

At this point, there now are what called “normal moments”, where the are now “normliazed” by dividing by the standard deviation. Now the  $i_{th}$  moment is given by

$$\frac{\mu_i}{\sigma^i}$$

The third normalized or standardized moment is called Pearson moment coefficient of skeweness.

$$M_3 = \gamma = \frac{\mu_3}{\sigma^3} = E\left[\left(\frac{s - \mu_1}{\sigma}\right)^3\right] = E\left[\frac{(E[(X - \mu)^3])}{E[(X - \mu)^2]^{\frac{3}{2}}}\right]$$

Confirming with r

```
x <- rgamma(200, shape = 10)
skewness(x)
```

```
## [1] 0.5260652
```

```
mean(mean((x - mean(x))^3)/(mean((x - mean(x))^2)^(3/2)))
```

```
## [1] 0.5260652
```

Forth Moment, kurtosis.

$$M_4 = \beta_2 = \frac{\mu_4}{\sigma^4} = E\left[\frac{(E[(X - \mu)^4])}{E[(X - \mu)^2]^2}\right]$$

Confirming with r

```
x <- rnorm(200)
kurtosis(x)
```

```
## [1] 2.761481
```

```
mean(mean((x - mean(x))^4)/(mean((x - mean(x))^2)^(2)))
```

```
## [1] 2.761481
```

However, because the kurtosis is equal to three in a normal distribution, is is often subtracted from the measure of kurtosis so that the sign of the value indicates whether a distribution is platykurtic (less than 0, low kurtosis) or leptokurtic (greater than 0, high kurtosis). It is simultaneously a measure of three values. It generally increases with peakedness, how much variance is right next to the mean and tails and the tail thickness, but decreases with increased mass held in the shoulders, or more modest deviations. However, it generally is seen as mostly a measure of tail thickness.

$$M_4 = \beta_2 = \frac{\mu_4}{\sigma^4} = E\left[\left(\frac{E[(X - \mu)^4]}{E[(X - \mu)^2]^2}\right) - 3\right]$$

## Central Limit theorem.

Proof by assertion. It's an interesting idea. What happens when you start to sum n, iid variables. This is interesting as it assumes its observation is itself a draw from an independent distribution (which I guess is restating its assumption). If you know the nature of the underlying distribution, you can calculate your expectation about the mean, variance, skew and kurtosis of the sum of the variables.

Rules for adding or summing variables.

$$\begin{aligned} \text{Mean}(X + Y) &= \text{Mean}(X) + \text{Mean}(Y) \\ \text{Mean}(aX) &= a * \text{Mean}(X) \\ \text{Var}(X + Y) &= \text{Var}(X) + \text{Var}(Y) \\ \text{Var}(aX) &= a^2 * \text{Var}(X) \end{aligned}$$

Question

$\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y)$  if  $Y = X$  why does  $\text{Var}(2*X) \neq \text{Var}(X + X)$  not clear, I guess different quantities

This means that if one starts to sum iid variables,  $n$  increases linearly. However, by assertion the kurtosis and skew start to fall towards zero, meaning that distribution of the sum of iid variables starts to look like a normal distribution. Apparently they increase like this, regardless of the initial distribution.

$$\text{skew}(\sum^n X) = \frac{1}{\sqrt{n}} \text{skew}(x)$$

$$\text{kurtosis}(\sum^n X) = \frac{1}{n} \text{kurtosis}(x)$$

This more or less underlies most statistics. There is the assumption that as you increase  $n$ , the expectation about the sum of the added random variables starts to resemble a normal variable (though normally you divide this value by  $n$ ). This is

## Normal variable

Here  $X$  is distributed according to the following function.

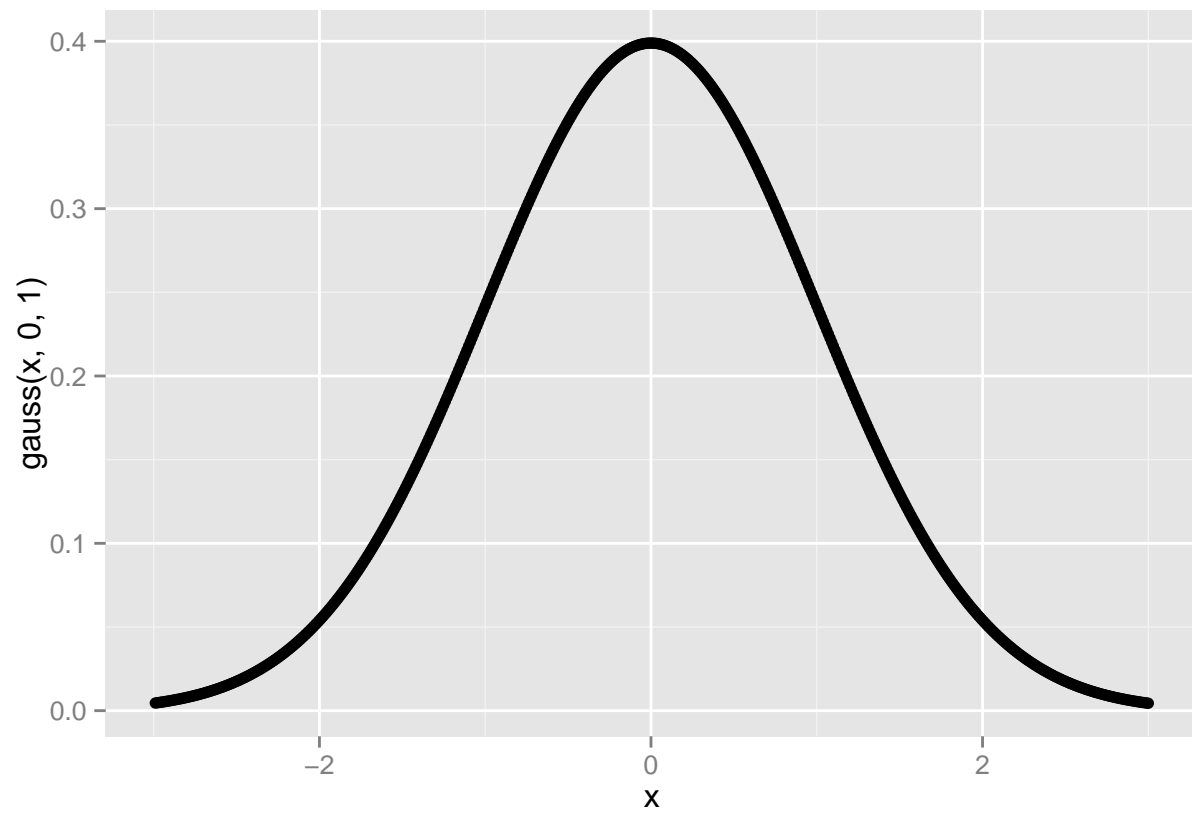
$$P(x|\mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\left[-\frac{x-\mu}{2\sigma^2}\right]}$$

Check in R

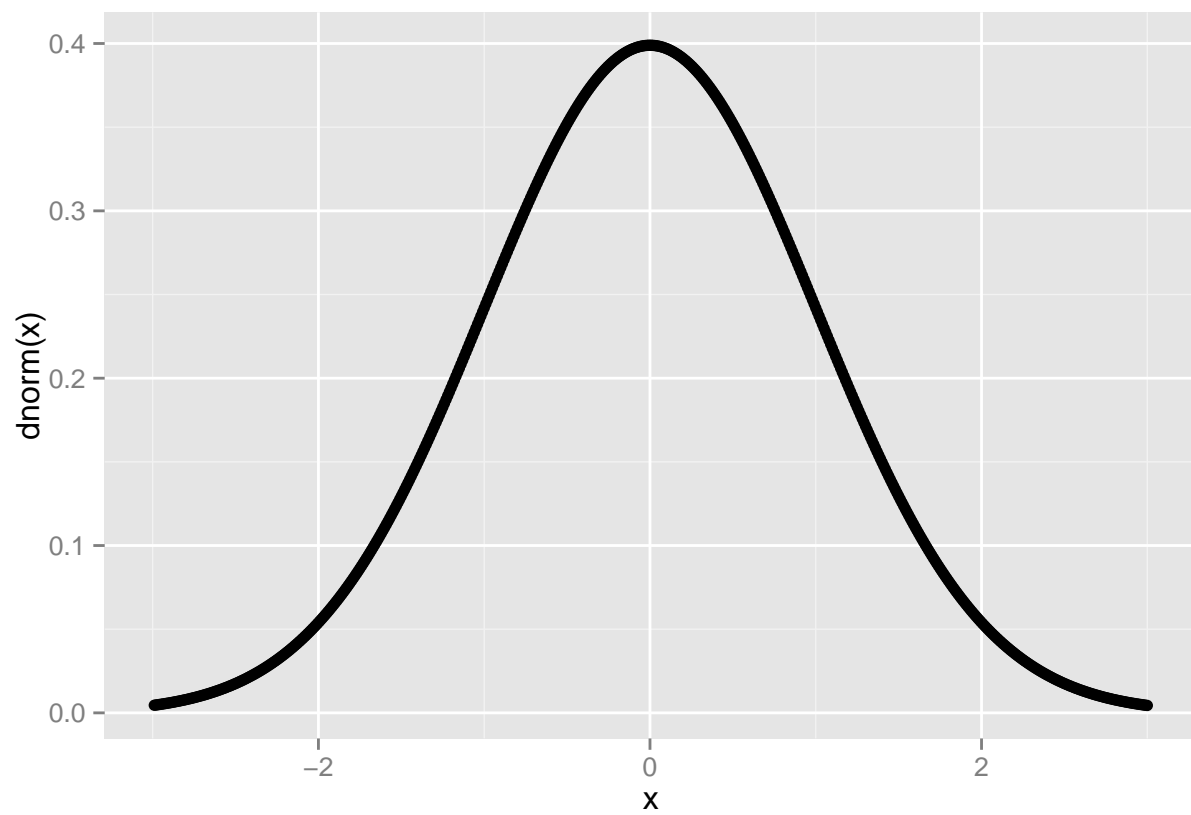
```
gauss <- function(x, mu, sigma){
  const <- 1 / sqrt( 2 * pi * sigma^2)
  probability <- const * exp( - (x-mu)^2 / (2 * sigma^2) )
  return(probability)
}

x<-1:600/100 - 3

qplot(x,gauss(x,0,1))
```



```
qplot(x,dnorm(x))
```



## **Law of large numbers**

If you have now taken the sum of this set of random variables, you can now get the expected value of the mean and variance by dividing by  $n$ . I think this means that there is a careful set

## **Multivariate extension**

## **Calculating Sample Variance**

## **Combining variances.**

Shape - skewness kurtosis

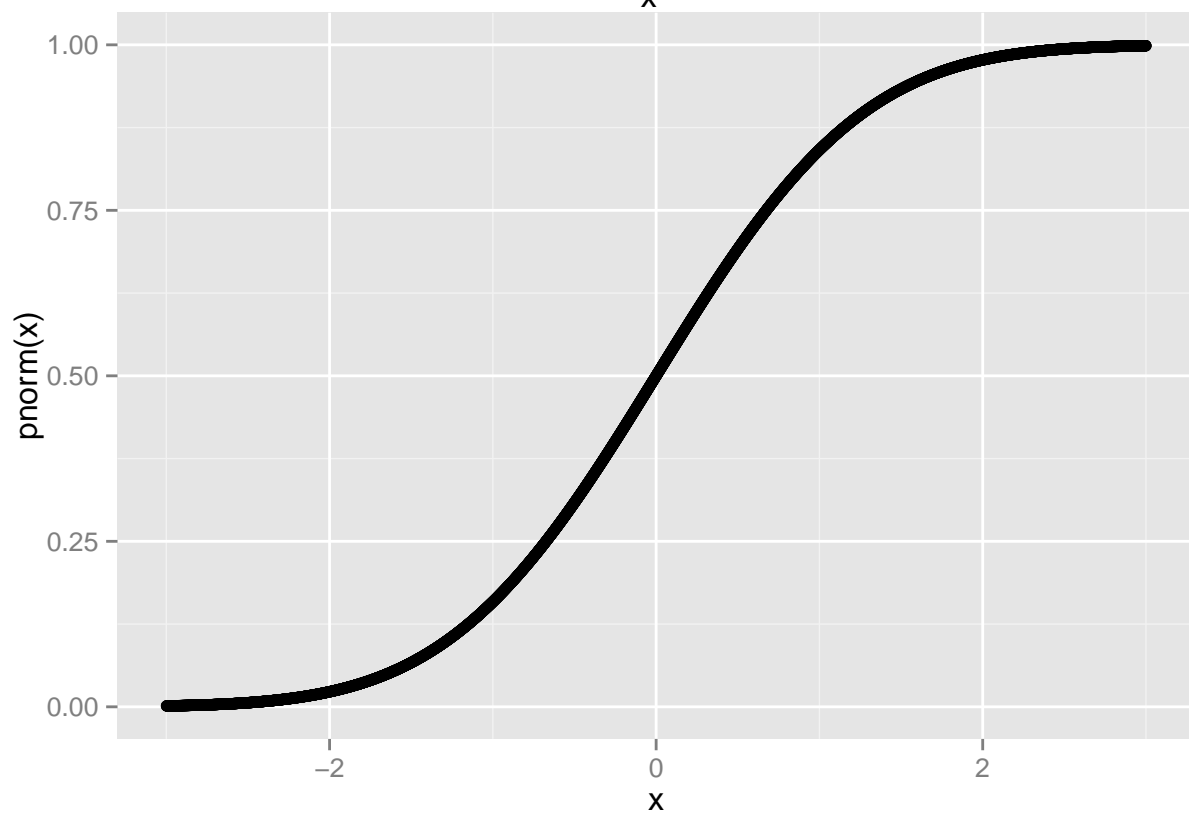
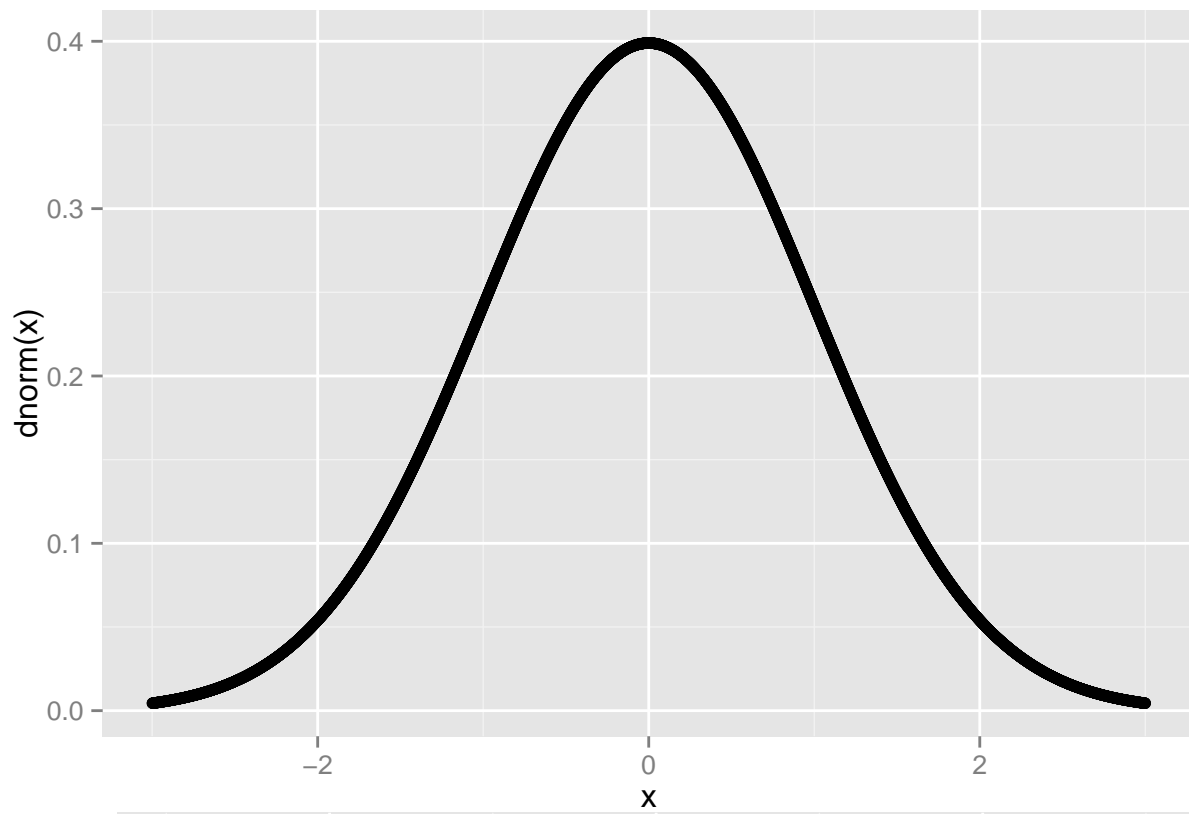
-How do you calculate these, can you recognize them in a graph.

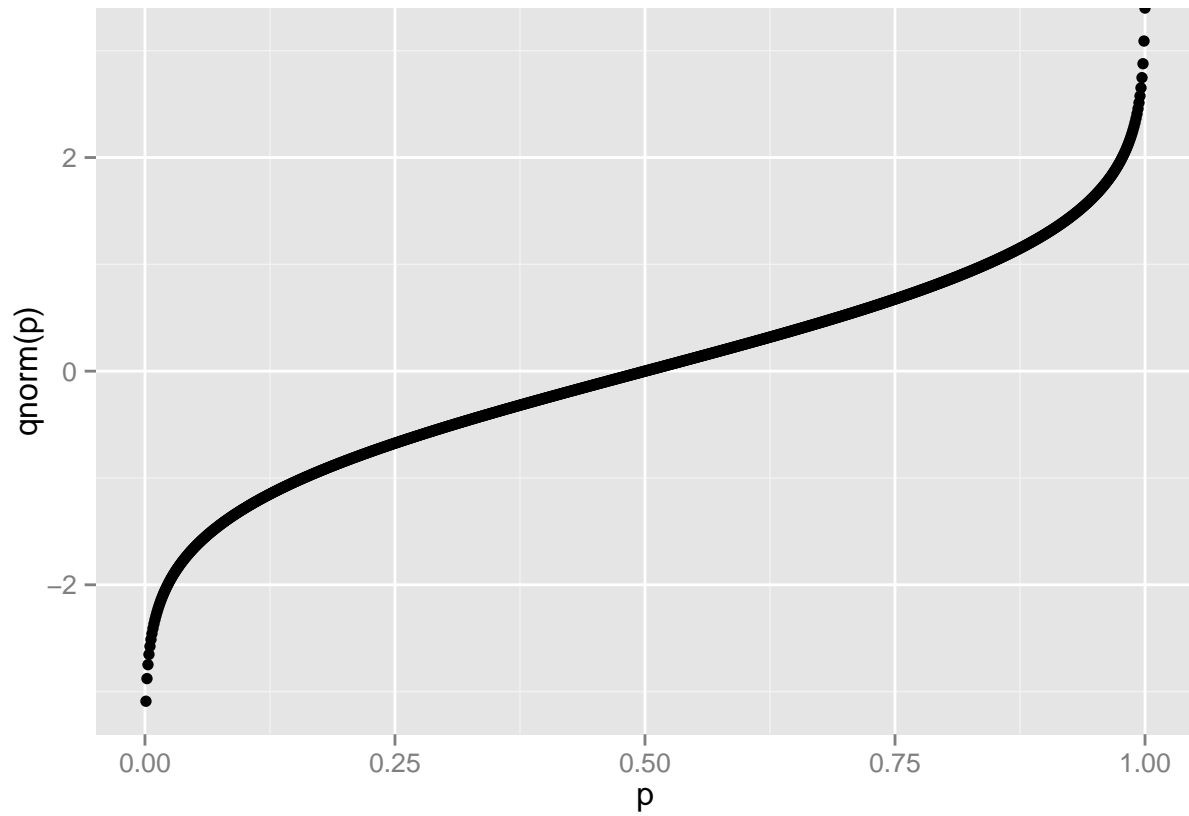
## **Reading and estimating from graphs**

## **Probability Distributions.**

## **pdf,cdf, icdf**

Here are the r commands, same for continuous and discrete probability density functions.





-Not refer to law of total probability in the answers, particularly as numerator in bayes equation.

## Random Variables and Expectation

### Null Hypothesis Sampling Testing.