Post-Graduate Statistics Course

Jarrod Hadfield

University of Edinburgh

Generalised Linear Mixed Models

- ANCOVA
- Bradley-Terry models
- MANCOVA
- Meta-analysis
- Multi-membership models
- Pedigree analysis: animal models
- Phylogenetic analysis: comparative approach
- Random Regression
- Rasch Models
- Regression
- Ridge Regression
- Splines
- Survival-analysis
- Threshold models
- Time-series

Course Outline

Course Outline

	Morning	Afternoon
Mon	The basics	
Tue		
Wed		
Thu		
Fri		

Course Outline

	Morning	Afternoon
Mon	The basics	
Tue	Linear Models	
Wed		
Thu		
Fri		

Course Outline

	Morning	Afternoon
Mon	The basics	
Tue	Linear Models	
Wed	Generalised Linear Models	
Thu		
Fri		

Course Outline

	Morning	Afternoon
Mon	The basics	
Tue	Linear Models	
Wed	Generalised Linear Models	
Thu	Mixed Models I	
Fri		

Course Outline

	Morning	Afternoon
Mon	The basics	
Tue	Linear Models	
Wed	Generalised Linear Models	
Thu	Mixed Models I	
Fri	Mixed Models II	

Course Outline

	Morning	Afternoon
Mon	The basics	Simulation
Tue	Linear Models	Linear Model Fitting in R
Wed	Generalised Linear Models	
Thu	Mixed Models I	Insights into Grumpiness
Fri	Mixed Models II	

Course Outline

	Morning	Afternoon
Mon	The basics	Simulation
Tue	Linear Models	Linear Model Fitting in R
Wed	Generalised Linear Models	
Thu	Mixed Models I	Insights into Grumpiness
Fri	Mixed Models II	Own data

The basics

- What do we want to learn from the data?

The basics

- What do we want to learn from the data?
- Ingredients (Model, Parameters, Data)

The basics

- What do we want to learn from the data?
- Ingredients (Model, Parameters, Data)
- Distributions (Data, Sampling, Posterior)

The basics

- What do we want to learn from the data?
- Ingredients (Model, Parameters, Data)
- Distributions (Data, Sampling, Posterior)
- Linear Model

What do we want to learn from the data?

What do we want to learn from the data?

1	2	3	4	5	6	7	8	9	10
Very Happy				OK					Very Grumpy

What do we want to learn from the data?

1	2	3	4	5	6	7	8	9	10
Very Happy				OK					Very Grumpy

- Test a hypothesis

What do we want to learn from the data?

1	2	3	4	5	6	7	8	9	10
Very Happy				OK					Very Grumpy

- Test a hypothesis
- Get a (quantitative) understanding

What do we want to learn from the data?

1	2	3	4	5	6	7	8	9	10
Very Happy				OK					Very Grumpy

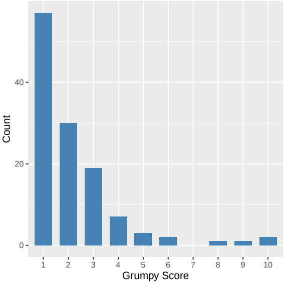
- Test a hypothesis
- Get a (quantitative) understanding
- Make a prediction

Ingredients

- Data
  - Response variable(s)
  - Predictor variable(s)

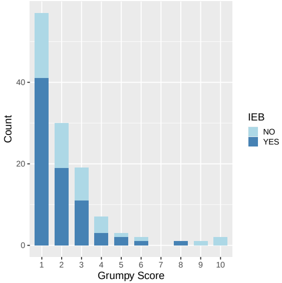
Ingredients

- Data
  - Response variable(s)
  - Predictor variable(s)



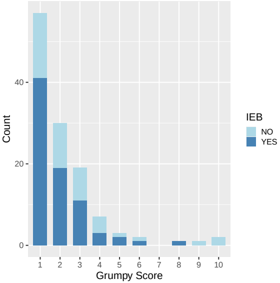
Ingredients

- Data
    - Response variable(s)
    - Predictor variable(s)

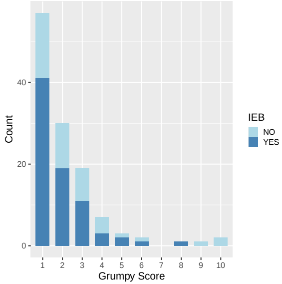


Ingredients

- Data
  - Response variable(s)
  - Predictor variable(s)
- Model
  - What distribution do the data follow?
  - How do the predictors change the data distribution?

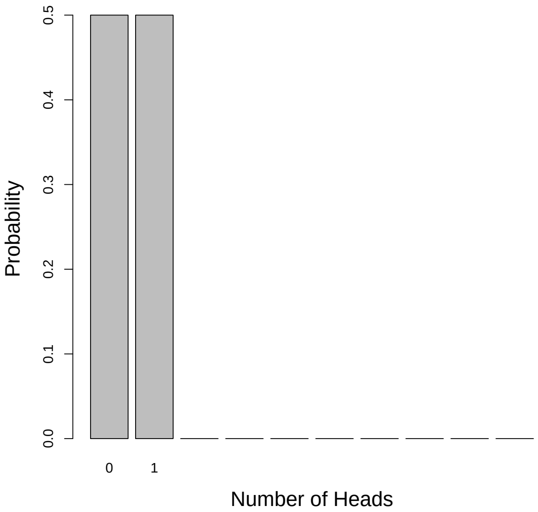


- Data
    - Response variable(s)
    - Predictor variable(s)
- Model
    - What distribution do the data follow?
    - How do the predictors change the data distribution?
- Parameters
    - Location
    - Dispersion

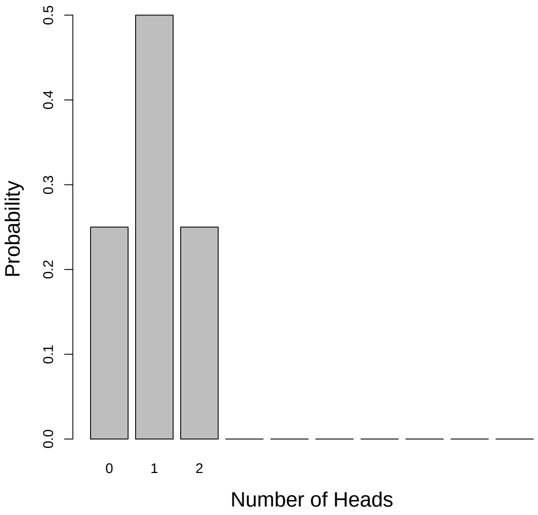


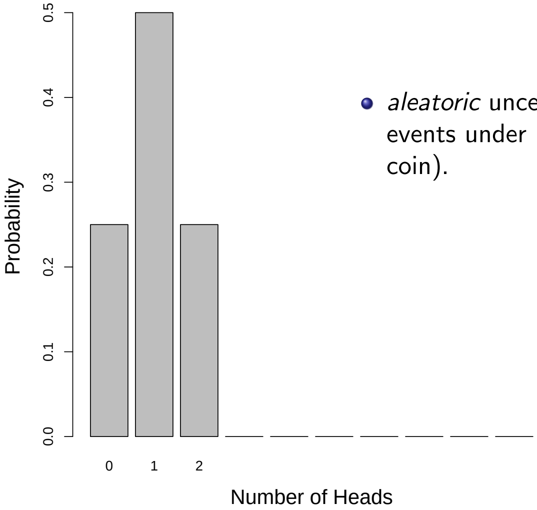
Probability Distributions

Probability Distributions



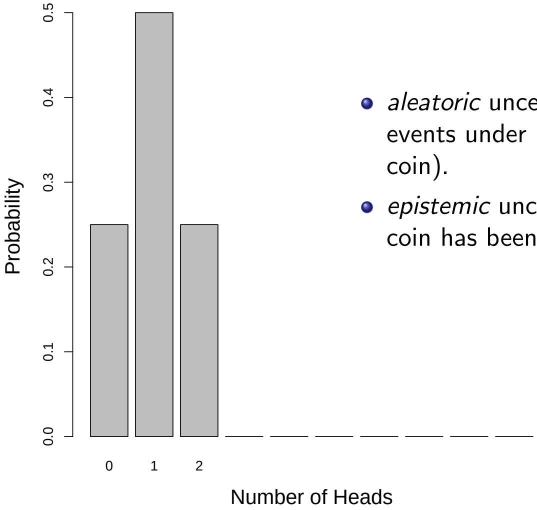
Probability Distributions





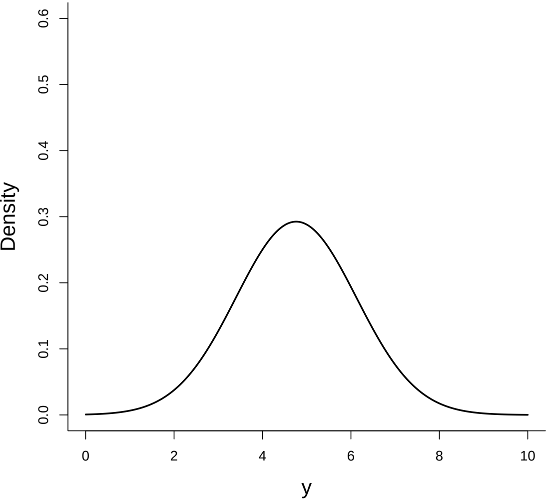
- *aleatoric* uncertainty is about the frequency of events under repeated sampling (tosses of the coin).

Probability Distributions

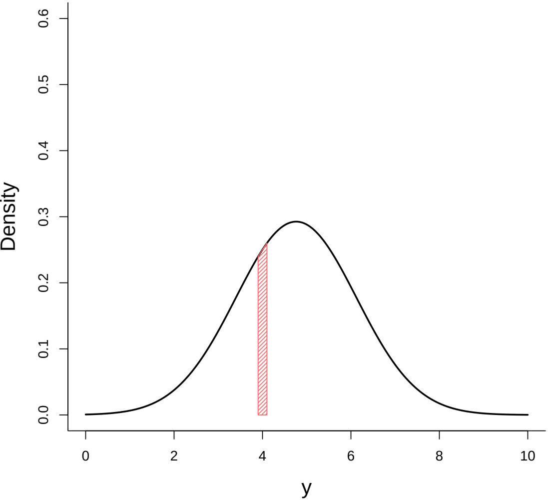


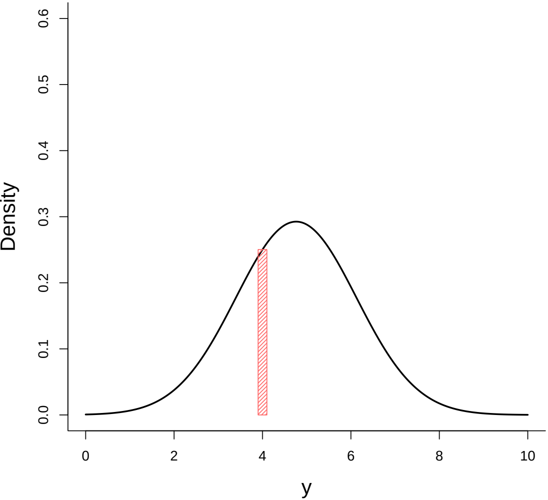
- *aleatoric* uncertainty is about the frequency of events under repeated sampling (tosses of the coin).
- *epistemic* uncertainty is about knowledge, the coin has been tossed.

Likelihood

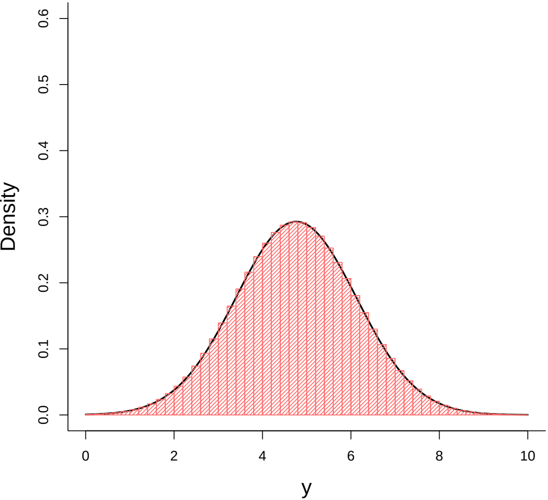


Likelihood

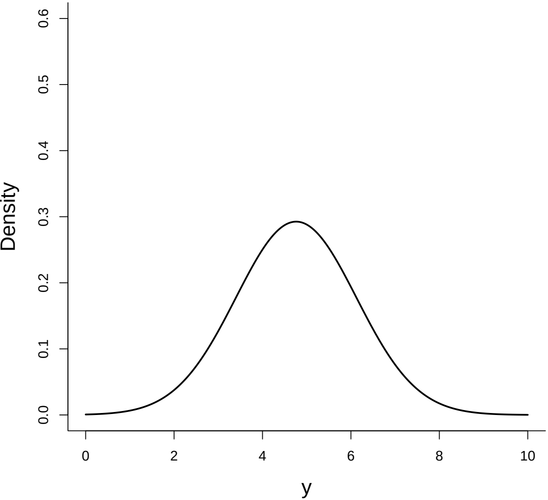



Density
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0
2
4
6
8
10
y

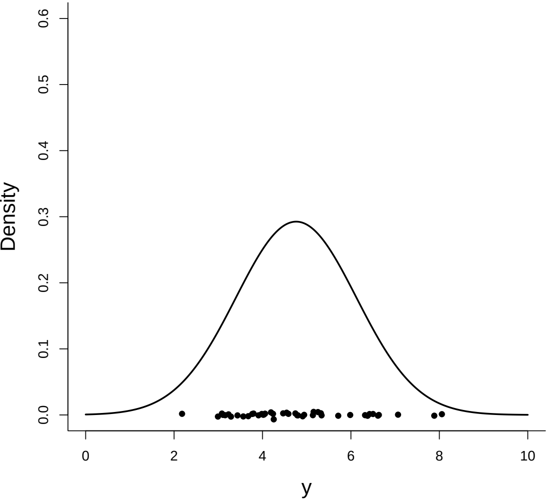
Likelihood



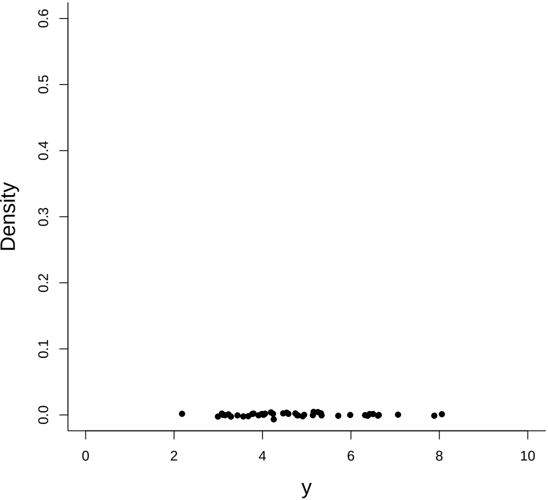
Likelihood



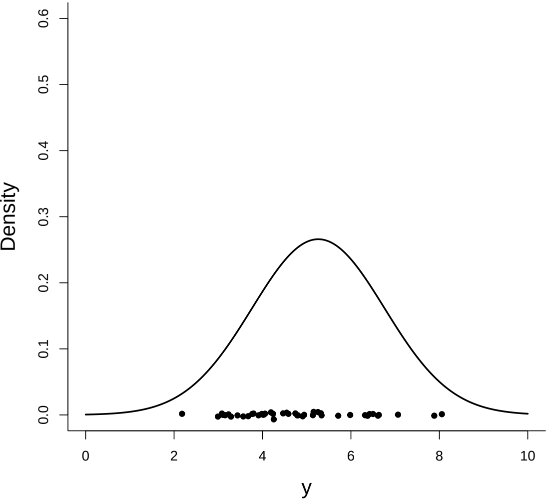
Likelihood



Likelihood



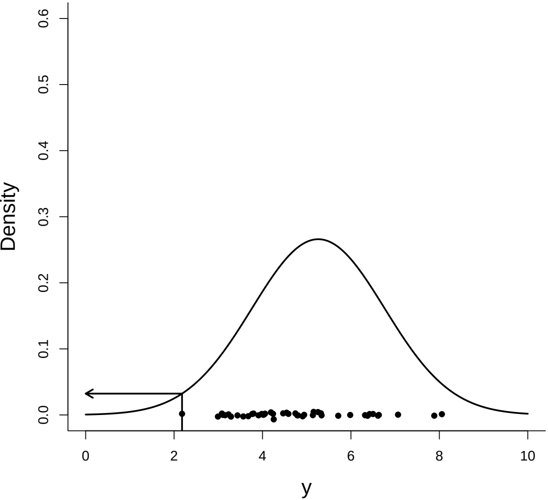
Likelihood



$$\widehat{\mu_D} = 5.3$$

$$\widehat{\sigma^2_D} = 2.3$$

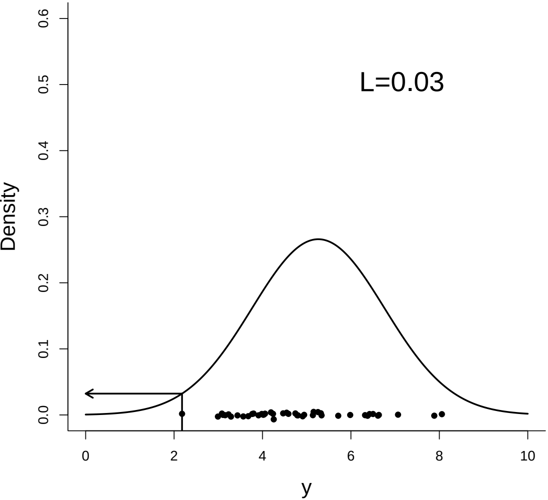
Likelihood



$$\widehat{\mu_D} = 5.3$$

$$\widehat{\sigma^2_D} = 2.3$$

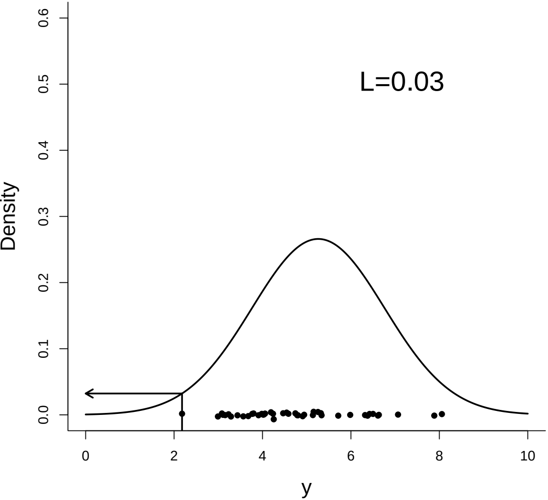
Likelihood



$$\widehat{\mu_D} = 5.3$$

$$\widehat{\sigma_D^2} = 2.3$$

Likelihood

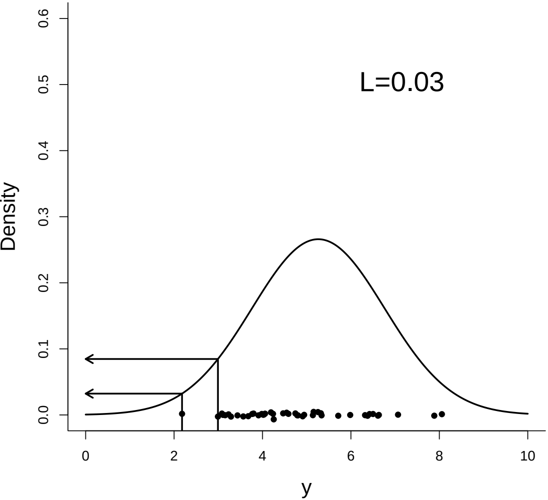


$\widehat{\mu_D} = 5.3$

$\widehat{\sigma^2_D} = 2.3$

```
> dnorm(2.2, 5.3, 1.5)
 0.03058795
```

Likelihood

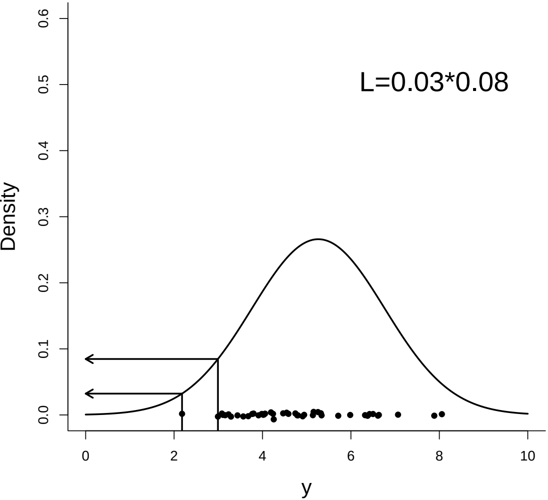


$\widehat{\mu_D} = 5.3$

$\widehat{\sigma^2_D} = 2.3$

```
> dnorm(3, 5.3, 1.5)
 0.0814022
```

Likelihood

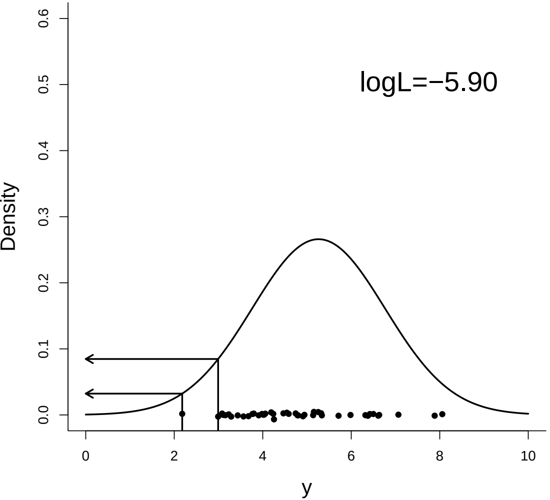


$\widehat{\mu_D} = 5.3$

$\widehat{\sigma^2_D} = 2.3$

```
> dnorm(3, 5.3, 1.5)
 0.0814022
```

Likelihood

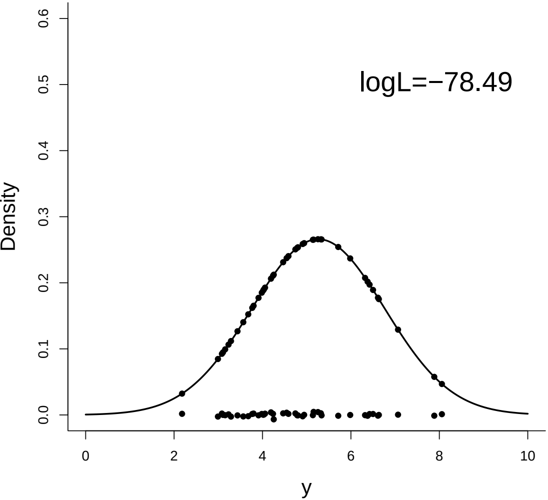


$$\widehat{\mu_D} = 5.3$$

$$\widehat{\sigma^2_D} = 2.3$$

```
> dnorm(3, 5.3, 1.5)
 0.0814022
```

Likelihood

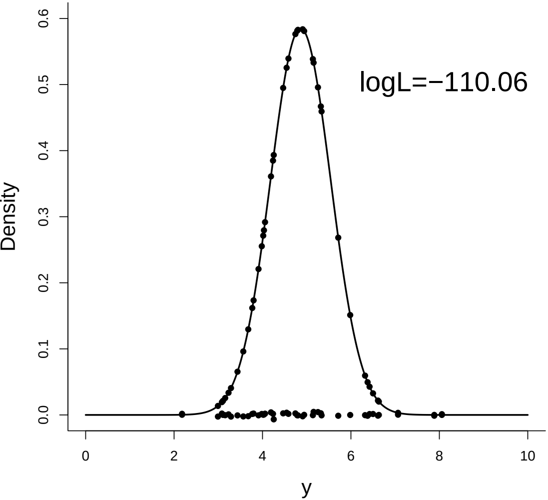


$\widehat{\mu_D} = 5.3$

$\widehat{\sigma^2_D} = 2.3$

```
> dnorm(3, 5.3, 1.5)
 0.0814022
```

Likelihood

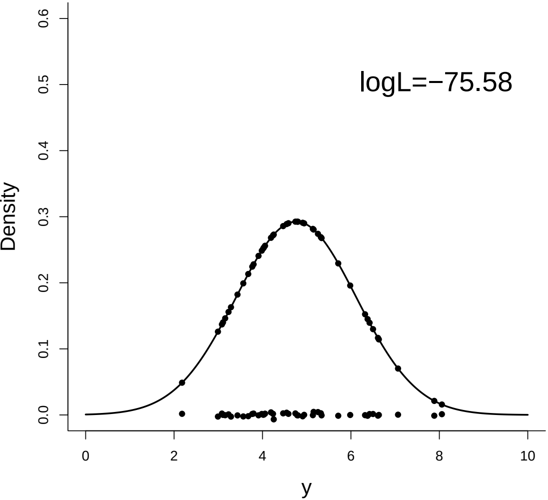


$\widehat{\mu_D} = 4.9$

$\widehat{\sigma^2_D} = 0.5$

```
> dnorm(3, 4.9, 0.7)
 0.0138733
```

Maximum Likelihood

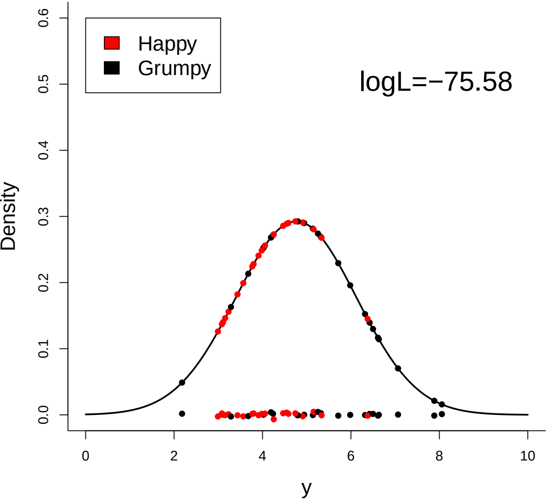


$\widehat{\mu_D} = 4.8$

$\widehat{\sigma^2_D} = 1.9$

```
> dnorm(3, 4.8, 1.4)
 0.12375071
```

Maximum Likelihood

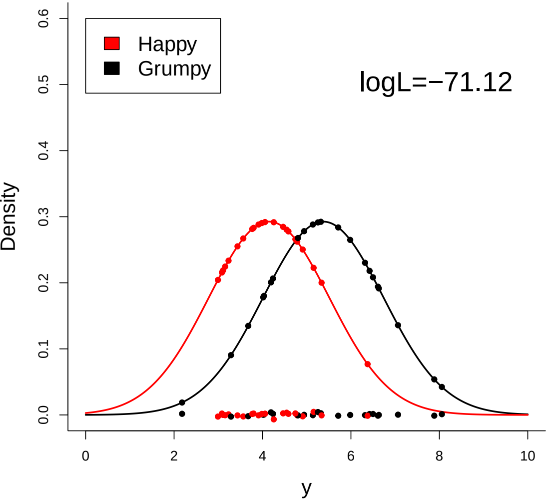


$\widehat{\mu_D} = 4.8$

$\widehat{\sigma^2_D} = 1.9$

```
> dnorm(3, 4.8, 1.4)
 0.12375071
```

Maximum Likelihood



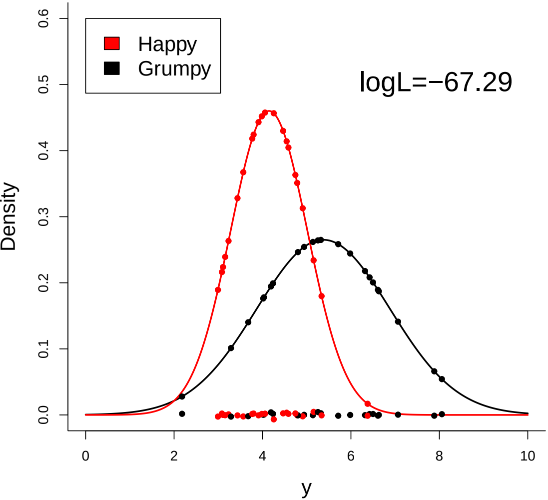
$\widehat{\mu_D} = 5.4$

$\widehat{\sigma^2_D} = 1.9$

```
> dnorm(3, 4.1, 1.4)
0.20831549
```

$\widehat{\mu_D} = 4.1$

$\widehat{\sigma^2_D} = 1.9$



$\widehat{\mu_D} = 5.4$

$\widehat{\sigma^2_D} = 2.3$

```
> dnorm(3, 4.1, 1.5)
  0.2076992
```

$\widehat{\mu_D} = 4.1$

$\widehat{\sigma^2_D} = 0.8$

Posterior Distribution

Likelihood: the *aleatoric* probability of the data *given* a parameter value.

Posterior Distribution

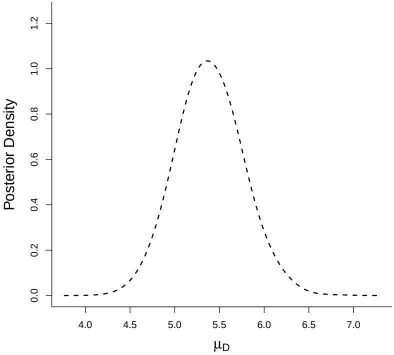
Likelihood: the *aleatoric* probability of the data *given* a parameter value.

Posterior Distribution: characterises *epistemic* uncertainty about the *true* parameter value.

Posterior Distribution

Likelihood: the *aleatoric* probability of the data *given* a parameter value.

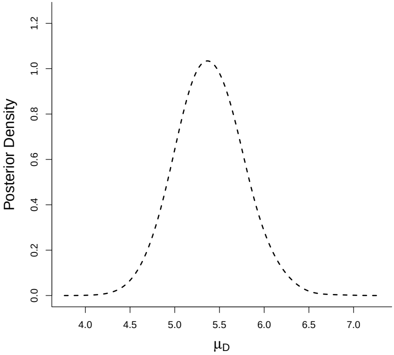
Posterior Distribution: characterises *epistemic* uncertainty about the *true* parameter value.



Posterior Distribution

Likelihood: the *aleatoric* probability of the data *given* a parameter value.

Posterior Distribution: characterises *epistemic* uncertainty about the *true* parameter value.



$$\text{Posterior} \propto \text{Likelihood} \times \text{Prior}$$

Sampling distribution: characterises *aleatoric* uncertainty about *estimates*.

Sampling Distribution

Sampling distribution: characterises *aleatoric* uncertainty about *estimates*.

Mind-bending ... but often similar to a posterior distribution.

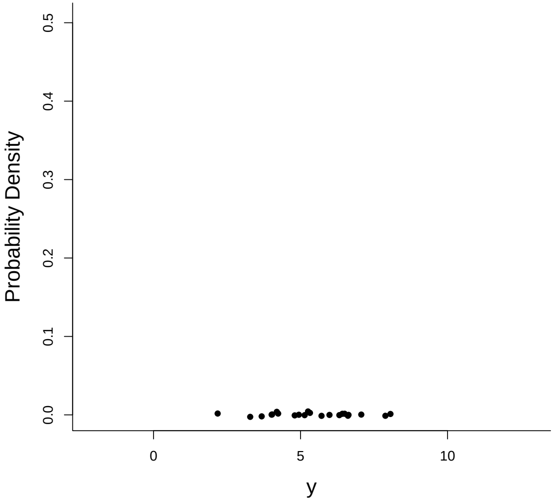
Sampling Distribution

Sampling distribution: characterises *aleatoric* uncertainty about *estimates*.

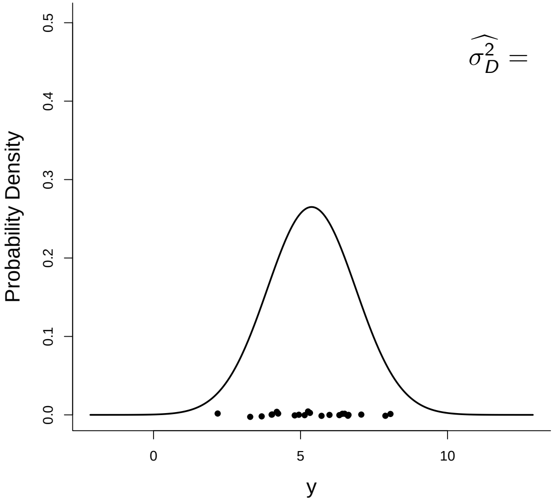
Mind-bending ... but often similar to a posterior distribution.

If you are a scientist rather than a statistician I want you to deceive yourselves that the sampling distribution is a posterior distribution but at the same time I want you to keep it in the back of your minds that you're being deceitful.

Sampling Distribution



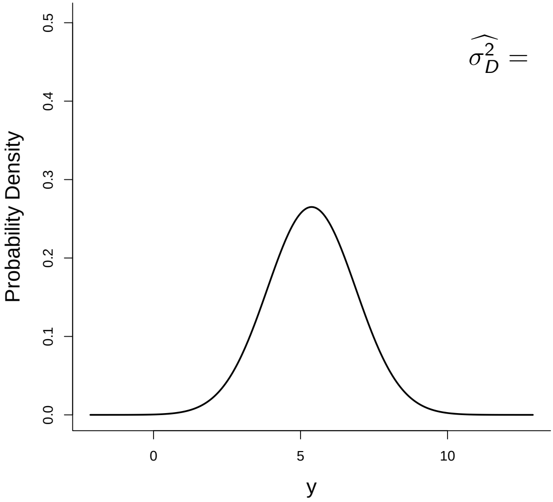
Sampling Distribution



$$\widehat{\mu_D} = \quad 5.38$$

$$\widehat{\sigma^2_D} = \quad 2.27$$

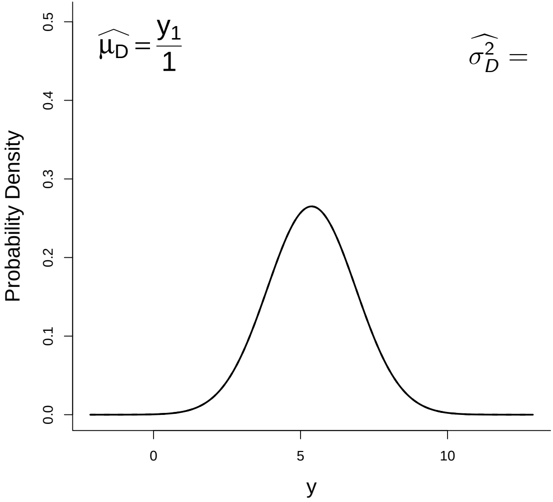
Sampling Distribution



$$\widehat{\mu_D} = \quad 5.38$$

$$\widehat{\sigma^2_D} = \quad 2.27$$

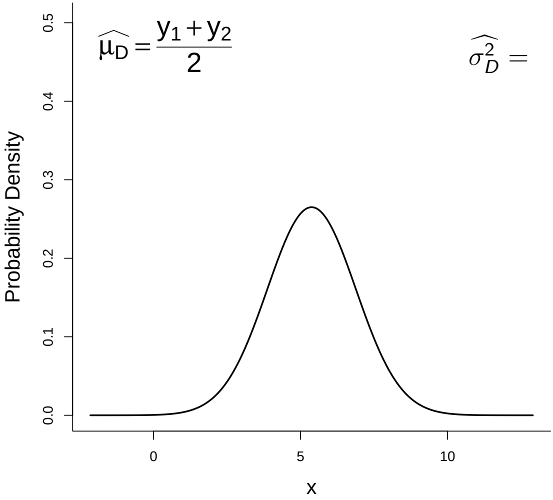
Sampling Distribution



$$\widehat{\mu_D} = 5.38$$

$$\widehat{\sigma^2_D} = 2.27$$

Sampling Distribution



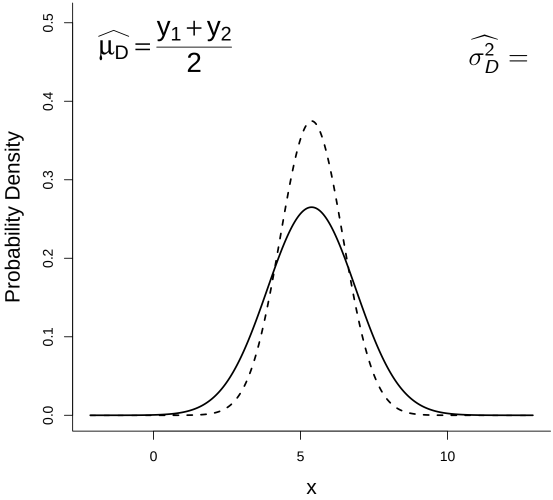
$$\widehat{\mu_D} = 5.38$$

$$\widehat{\sigma^2_D} = 2.27$$

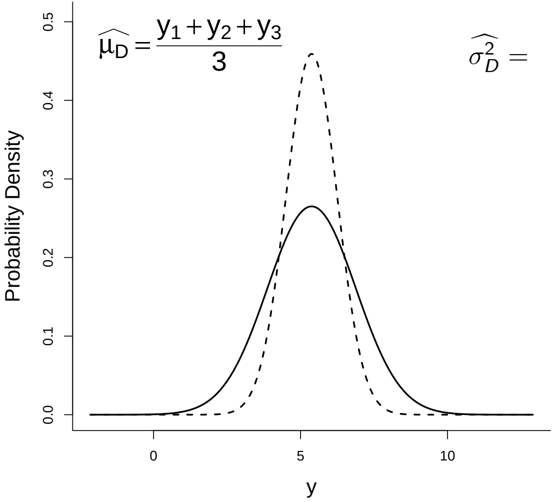
Sampling Distribution

$\widehat{\mu_D} = \quad 5.38$

$\widehat{\sigma^2_D} = \quad 2.27$



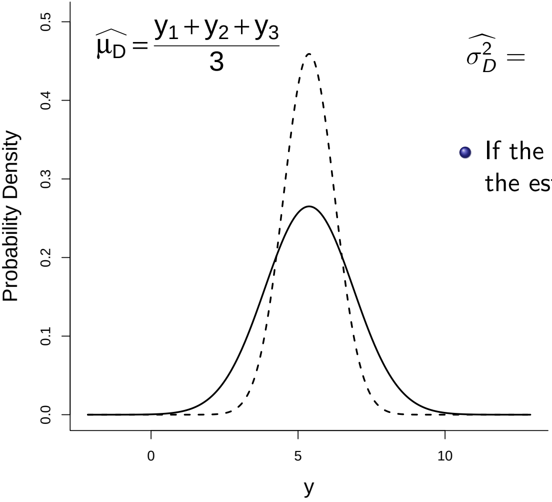
Sampling Distribution



$$\widehat{\mu_D} = 5.38$$

$$\widehat{\sigma^2_D} = 2.27$$

Sampling Distribution



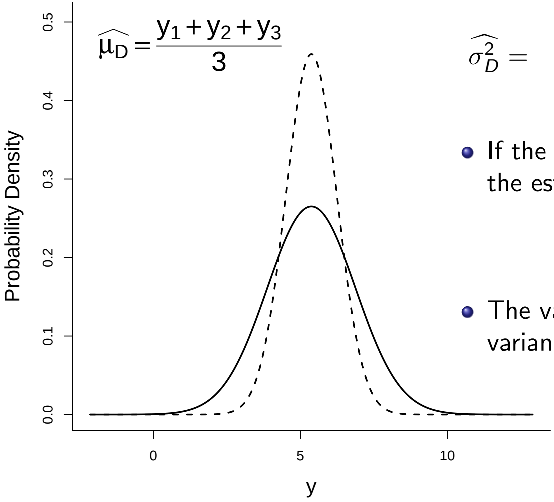
$$\widehat{\mu_D} = \quad 5.38$$

$$\widehat{\sigma^2_D} = \quad 2.27$$

- If the estimator is unbiased the mean of the estimator is equal to the true value:

$$E[\widehat{\mu_D}] = \mu_D$$

Sampling Distribution



$$\widehat{\mu_D} = \quad 5.38$$

$$\widehat{\sigma^2_D} = \quad 2.27$$

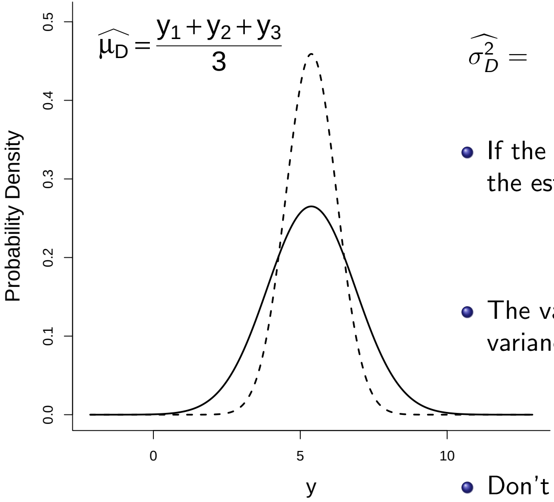
- If the estimator is unbiased the mean of the estimator is equal to the true value:

$$E[\widehat{\mu_D}] = \mu_D$$

- The variance of the estimator is the data variance divided by sample size:

$$\text{Var}[\widehat{\mu_D}] = \frac{\sigma^2_D}{n}$$

Sampling Distribution

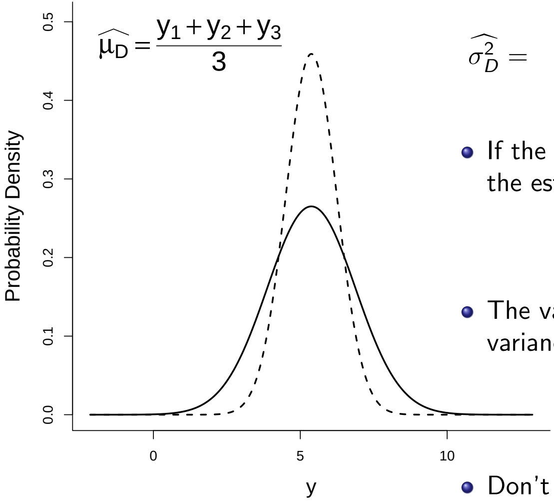


$$\widehat{\mu_D} = \quad 5.38$$

$$\widehat{\sigma^2_D} = \quad 2.27$$

- If the estimator is unbiased the mean of the estimator is equal to the true value:
$$E[\widehat{\mu_D}] = \mu_D$$
- The variance of the estimator is the data variance divided by sample size:
$$\text{Var}[\widehat{\mu_D}] = \frac{\sigma^2_D}{n}$$
- Don't know σ^2_D (use $\widehat{\sigma^2_D}$ instead):
$$\text{Var}[\widehat{\mu_D}] \approx \frac{\widehat{\sigma^2_D}}{n}$$

Sampling Distribution

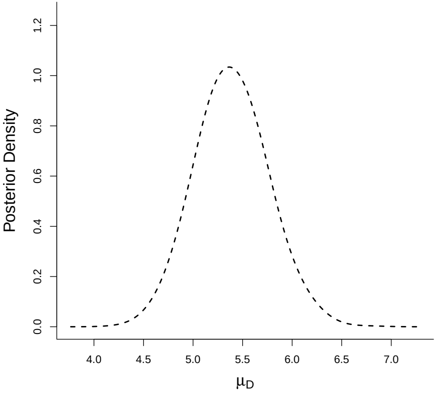


$$\widehat{\mu_D} = \quad 5.38$$

$$\widehat{\sigma^2_D} = \quad 2.27$$

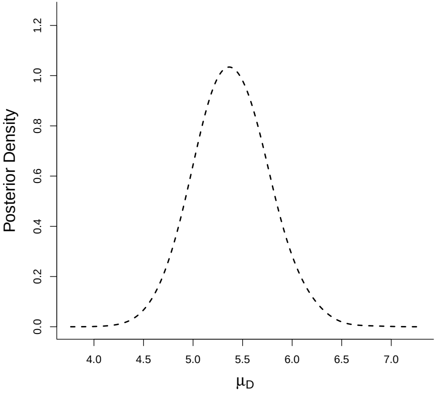
- If the estimator is unbiased the mean of the estimator is equal to the true value:
$$E[\widehat{\mu_D}] = \mu_D$$
- The variance of the estimator is the data variance divided by sample size:
$$\text{Var}[\widehat{\mu_D}] = \frac{\sigma^2_D}{n}$$
- Don't know σ^2_D (use $\widehat{\sigma^2_D}$ instead):
$$\text{Var}[\widehat{\mu_D}] \approx \frac{\widehat{\sigma^2_D}}{n} = \frac{2.27}{22} = 0.10$$

Sampling Distribution versus Posterior Distribution



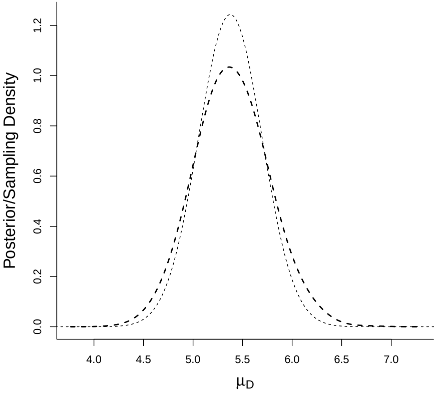
Sampling Distribution versus Posterior Distribution

As the information content of the data (relative to the prior) increases the sampling distribution tends to the posterior distribution.



Sampling Distribution versus Posterior Distribution

As the information content of the data (relative to the prior) increases the sampling distribution tends to the posterior distribution.

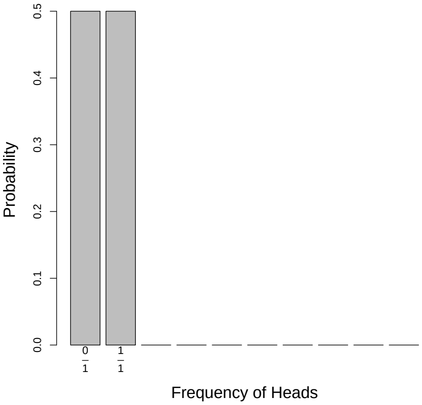


Sampling Distribution tends to a Normal

As the information content of the data increases the sampling distribution tends to a normal distribution (even if the data distribution is not normal)

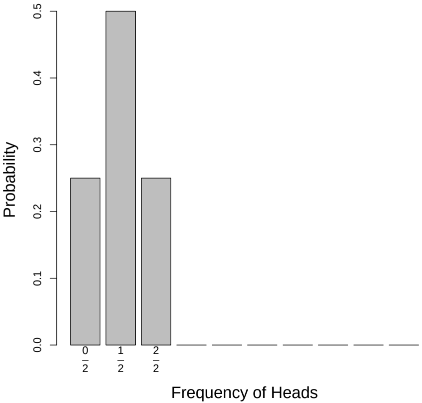
Sampling Distribution tends to a Normal

As the information content of the data increases the sampling distribution tends to a normal distribution (even if the data distribution is not normal)



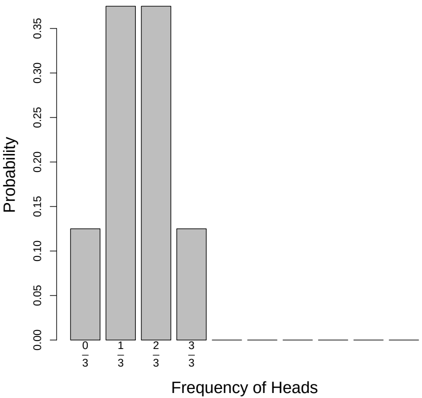
Sampling Distribution tends to a Normal

As the information content of the data increases the sampling distribution tends to a normal distribution (even if the data distribution is not normal)



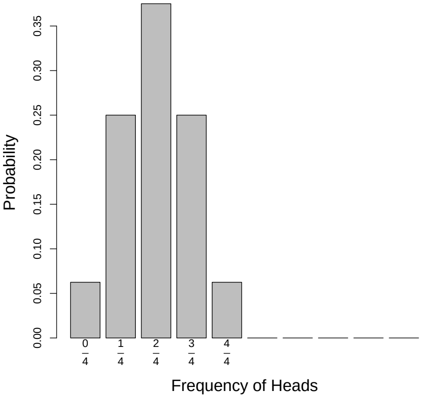
Sampling Distribution tends to a Normal

As the information content of the data increases the sampling distribution tends to a normal distribution (even if the data distribution is not normal)



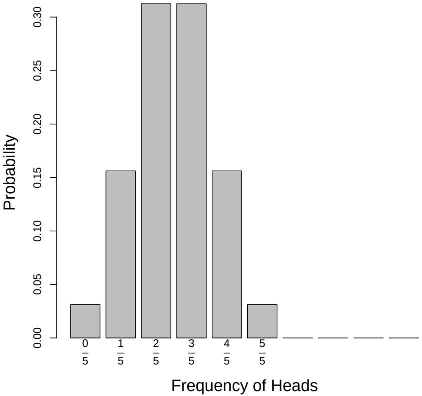
Sampling Distribution tends to a Normal

As the information content of the data increases the sampling distribution tends to a normal distribution (even if the data distribution is not normal)



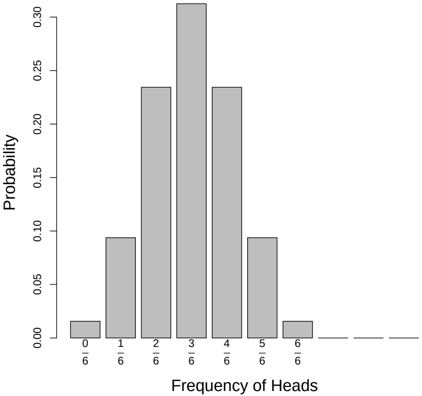
Sampling Distribution tends to a Normal

As the information content of the data increases the sampling distribution tends to a normal distribution (even if the data distribution is not normal)



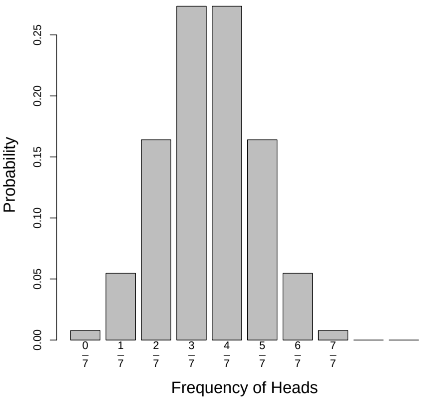
Sampling Distribution tends to a Normal

As the information content of the data increases the sampling distribution tends to a normal distribution (even if the data distribution is not normal)



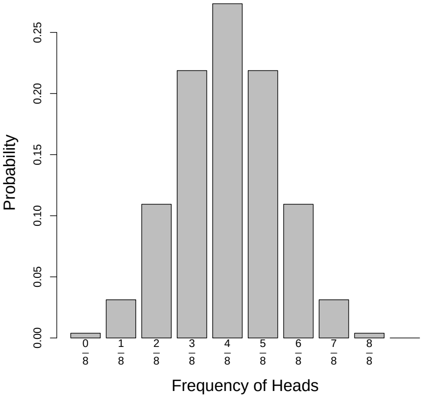
Sampling Distribution tends to a Normal

As the information content of the data increases the sampling distribution tends to a normal distribution (even if the data distribution is not normal)



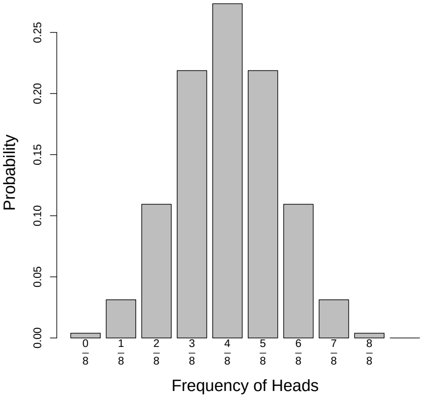
Sampling Distribution tends to a Normal

As the information content of the data increases the sampling distribution tends to a normal distribution (even if the data distribution is not normal)



Sampling Distribution tends to a Normal

As the information content of the data increases the sampling distribution tends to a normal distribution (even if the data distribution is not normal)



Thanks to the central limit theorem

Uncertainty and Distributions: mini-quiz

- Uncertainty

- Uncertainty

A rabid birdwatcher is chopping the heads off half the cats they meet.

- Uncertainty

A rabid birdwatcher is chopping the heads off half the cats they meet.

what is the probability a cat they meet will die?

what is the probability the cat they met is dead?

Uncertainty and Distributions: mini-quiz

- Uncertainty

A rabid birdwatcher is chopping the heads off half the cats they meet.

what is the probability a cat they meet will die?

what is the probability the cat they met is dead?

- Distributions

Uncertainty and Distributions: mini-quiz

- Uncertainty

A rabid birdwatcher is chopping the heads off half the cats they meet.

what is the probability a cat they meet will die?

what is the probability the cat they met is dead?

- Distributions

They've given 10 cats polonium, 8 are dead.

- Uncertainty

A rabid birdwatcher is chopping the heads off half the cats they meet.

what is the probability a cat they meet will die?

what is the probability the cat they met is dead?

- Distributions

They've given 10 cats polonium, 8 are dead.

what is the probability (p) a cat will die after polonium?

The home office has let them redo the experiment a gazzilian time
In how many experiments did 3 out of 10 cats die?

They have 10 trials and 'success' with probability p.

Uncertainty and Distributions: mini-quiz

- Uncertainty

A rabid birdwatcher is chopping the heads off half the cats they meet.

aleatoric: what is the probability a cat they meet will die?

epistemic: what is the probability the cat they met is dead?

- Distributions

They've given 10 cats polonium, 8 are dead.

what is the probability (p) a cat will die after polonium?

The home office has let them redo the experiment a gazzilian time
In how many experiments did 3 out of 10 cats die?

They have 10 trials and 'success' with probability p.

Uncertainty and Distributions: mini-quiz

- Uncertainty

A rabid birdwatcher is chopping the heads off half the cats they meet.

aleatoric: what is the probability a cat they meet will die?

epistemic: what is the probability the cat they met is dead?

- Distributions

They've given 10 cats polonium, 8 are dead.

posterior: what is the probability (p) a cat will die after polonium?

sampling: The home office has let them redo the experiment a gazzilian time
In how many experiments did 3 out of 10 cats die?

data: They have 10 trials and 'success' with probability p.

lm

$\widehat{\mu_D} = 5.376 \quad \widehat{\sigma^2_D} = 2.266 \quad \text{Var}[\widehat{\mu_D}] = 0.103$

lm

$$\widehat{\mu_D} = 5.376 \quad \widehat{\sigma^2_D} = 2.266 \quad \text{Var}[\widehat{\mu_D}] = 0.103$$

$$\widehat{\sigma_D} = 1.505 \quad \text{SD}[\widehat{\mu_D}] = 0.321$$

lm

$$\widehat{\mu_D} = 5.376 \quad \widehat{\sigma^2_D} = 2.266 \quad \text{Var}[\widehat{\mu_D}] = 0.103$$

$$\widehat{\sigma_D} = 1.505 \quad \text{SD}[\widehat{\mu_D}] = 0.321$$

```
> photo_m1 <- lm(y ~ 1, data = subset(photo_long,
+     type == "grumpy"))
```

$$\widehat{\mu_D} = 5.376 \quad \widehat{\sigma^2_D} = 2.266 \quad \text{Var}[\widehat{\mu_D}] = 0.103$$

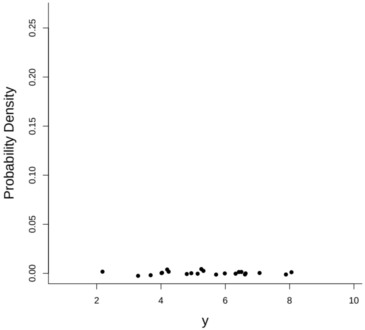
$$\widehat{\sigma_D} = 1.505 \quad \text{SD}[\widehat{\mu_D}] = 0.321$$

```
> photo_m1 <- lm(y ~ 1, data = subset(photo_long,
+      type == "grumpy"))
> summary(photo_m1)
Residuals:
   Min     1Q Median     3Q    Max
-3.196 -1.173 -0.089  1.105  2.682

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   5.3759     0.3209   16.75 1.25e-13 ***
---
Signif. codes:
0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.505 on 21 degrees of freedom
```

lm: Residuals

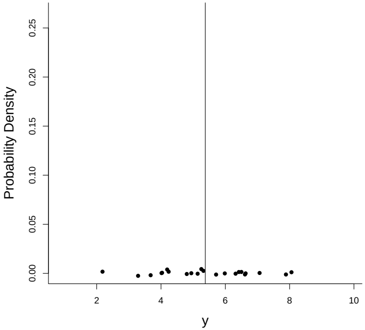
```
> summary(photo_m1)
Residuals:
   Min     1Q Median     3Q    Max
-3.196 -1.173 -0.089  1.105  2.682
```

```
> summary(photo_m1)
Residuals:
   Min     1Q Median     3Q    Max
-3.196 -1.173 -0.089  1.105  2.682
```

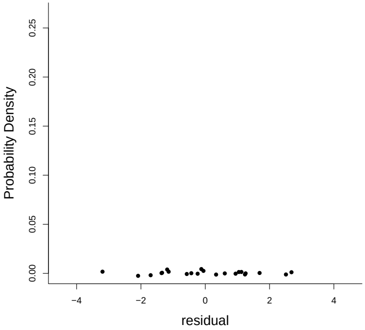


lm: Residuals

```
> summary(photo_m1)
Residuals:
   Min     1Q Median     3Q    Max
-3.196 -1.173 -0.089  1.105  2.682
```

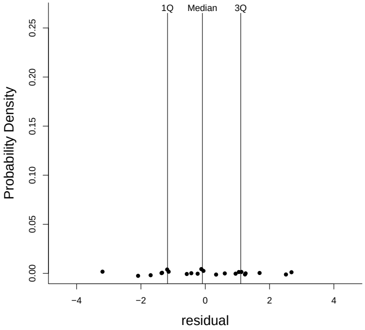


```
> summary(photo_m1)
Residuals:
   Min     1Q Median     3Q    Max
-3.196 -1.173 -0.089  1.105  2.682
```



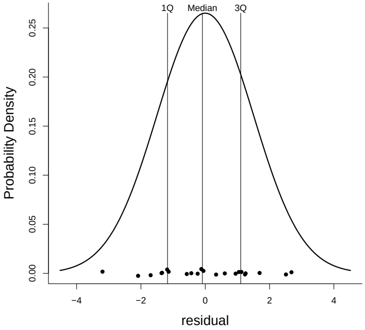
lm: Residuals

```
> summary(photo_m1)
Residuals:
   Min     1Q  Median     3Q    Max
-3.196 -1.173 -0.089  1.105  2.682
```



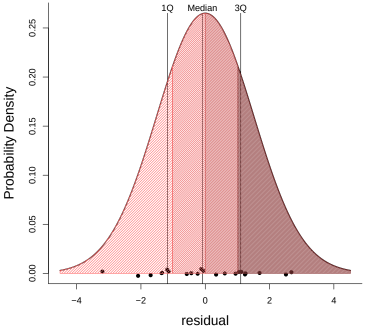
lm: Residuals

```
> summary(photo_m1)
Residuals:
   Min     1Q Median     3Q    Max
-3.196 -1.173 -0.089  1.105  2.682
```

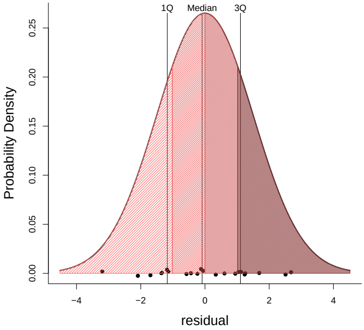


lm: Residuals

```
> summary(photo_m1)
Residuals:
   Min     1Q  Median     3Q    Max
-3.196 -1.173 -0.089  1.105  2.682
```



```
> summary(photo_m1)
Residuals:
   Min     1Q  Median     3Q    Max
-3.196 -1.173 -0.089  1.105  2.682
```

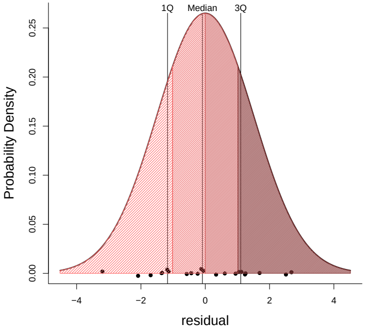


Quantile Function

```
> qnorm(1/4, 0, sd.hat)
[1] -1.015226
```

lm: Residuals

```
> summary(photo_m1)
Residuals:
    Min      1Q  Median      3Q     Max
-3.196 -1.173 -0.089  1.105  2.682
```



Quantile Function

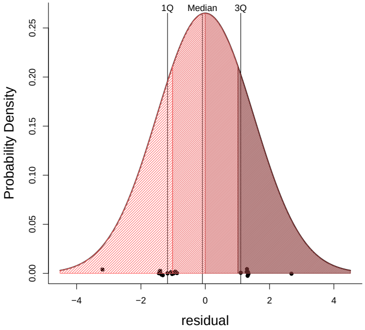
```
> qnorm(1/4, 0, sd.hat)
[1] -1.015226
```

Cumulative Distribution Function

```
> pnorm(-1.05226, 0, sd.hat)
[1] 0.25
```

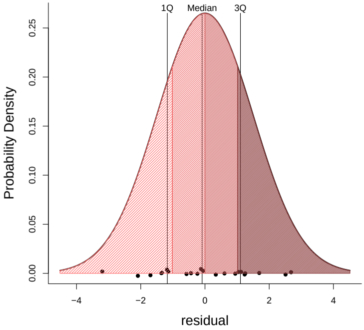
lm: Residuals

```
> summary(photo_m1)
Residuals:
    Min      1Q  Median      3Q     Max
-3.196 -1.173 -0.089  1.105  2.682
```

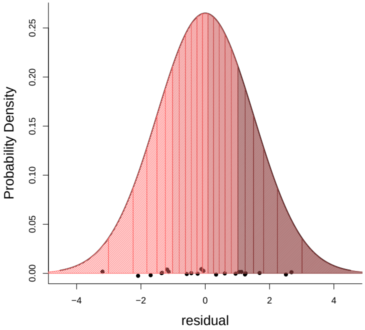


lm: Residuals

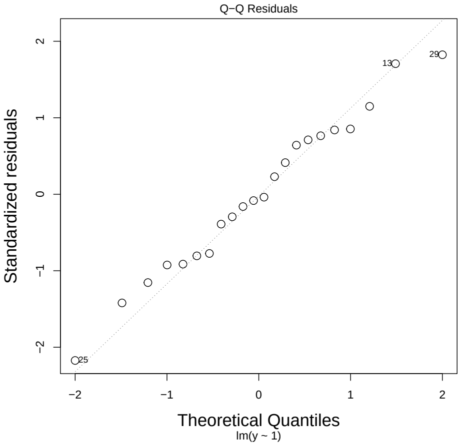
```
> summary(photo_m1)
Residuals:
    Min      1Q  Median      3Q     Max
-3.196 -1.173 -0.089   1.105   2.682
```



```
> summary(photo_m1)
Residuals:
   Min     1Q Median     3Q    Max
-3.196 -1.173 -0.089  1.105  2.682
```



```
> plot(photo_m1, 2)
```

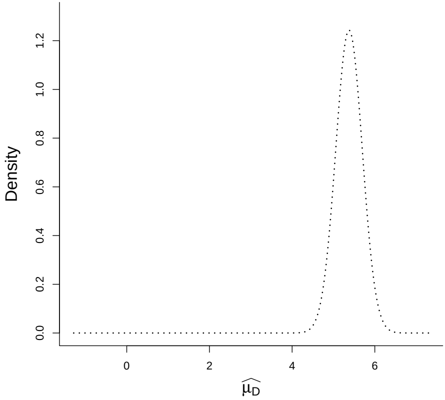


`lm`: Confidence Intervals

```
> coef(summary(photo_m1))
            Estimate Std. Error  t value     Pr(>|t|)
(Intercept) 5.375865  0.3209046 16.75222 1.254999e-13
```

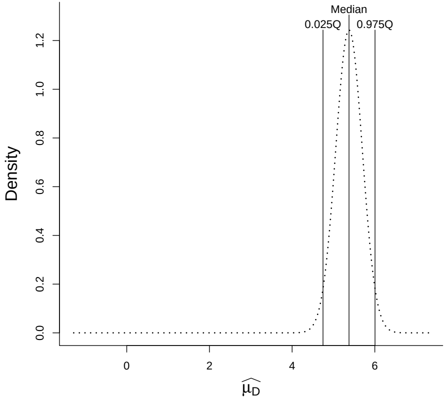
lm: Confidence Intervals

```
> coef(summary(photo_m1))
             Estimate Std. Error  t value     Pr(>|t|)
(Intercept) 5.375865  0.3209046 16.75222 1.254999e-13
```

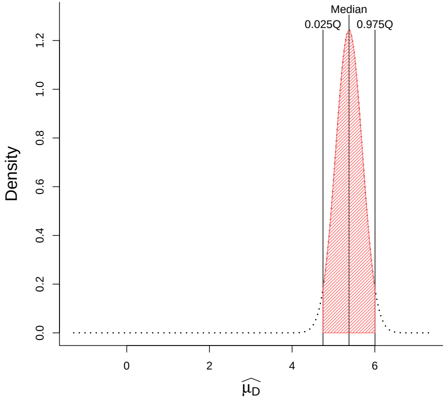


lm: Confidence Intervals

```
> coef(summary(photo_m1))
            Estimate Std. Error  t value     Pr(>|t|)
(Intercept) 5.375865  0.3209046 16.75222 1.254999e-13
```

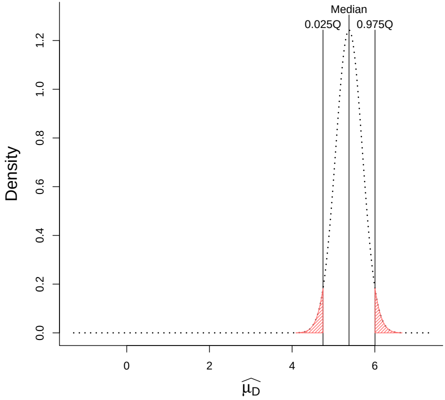


```
> coef(summary(photo_m1))
            Estimate Std. Error  t value     Pr(>|t|)
(Intercept) 5.375865  0.3209046 16.75222 1.254999e-13
```



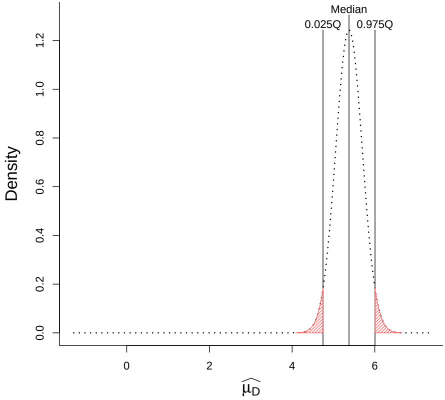
lm: Confidence Intervals

```
> coef(summary(photo_m1))
            Estimate Std. Error  t value     Pr(>|t|)
(Intercept) 5.375865  0.3209046 16.75222 1.254999e-13
```



lm: Confidence Intervals

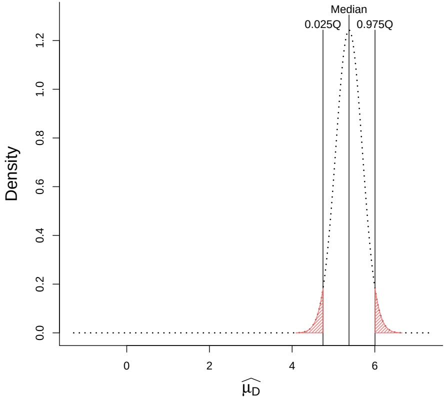
```
> coef(summary(photo_m1))
            Estimate Std. Error  t value     Pr(>|t|)
(Intercept) 5.375865  0.3209046 16.75222 1.254999e-13
```




```
> confint(photo_m1)
               2.5 %   97.5 %
(Intercept) 4.708507 6.043222
```

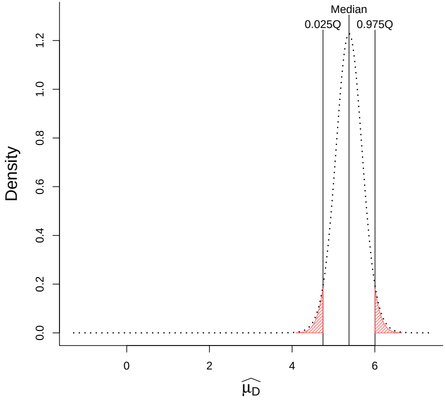
lm: Confidence Intervals

```
> coef(summary(photo_m1))
             Estimate Std. Error  t value     Pr(>|t|)
(Intercept) 5.375865  0.3209046 16.75222 1.254999e-13
```



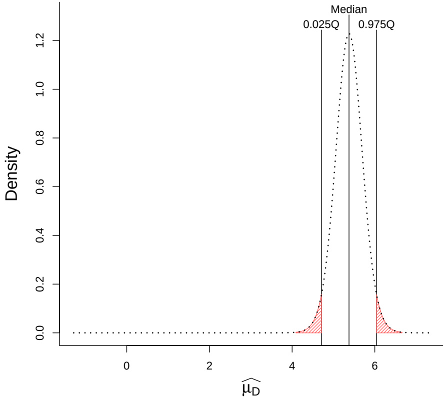

```
> confint(photo_m1)
               2.5 %   97.5 %
(Intercept) 4.708507 6.043222

> qnorm(c(0.025, 0.975),
+      mean = mu.hat, sd = mu.se)
[1] 4.746904 6.004826
```

```
> coef(summary(photo_m1))
            Estimate Std. Error  t value     Pr(>|t|)
(Intercept) 5.375865  0.3209046 16.75222 1.254999e-13
```




```
> confint(photo_m1)
               2.5 %   97.5 %
(Intercept) 4.708507 6.043222

> qnorm(c(0.025, 0.975),
+       mean = mu.hat, sd = mu.se)
[1] 4.746904 6.004826
```

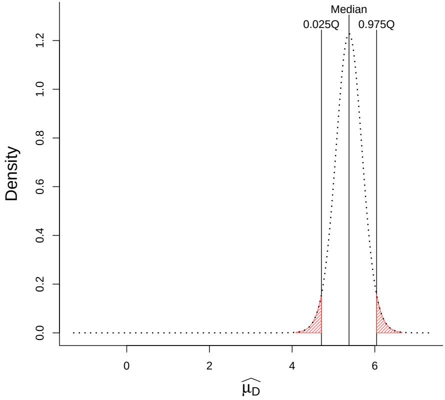
```
> coef(summary(photo_m1))
             Estimate Std. Error  t value     Pr(>|t|)
(Intercept) 5.375865  0.3209046 16.75222 1.254999e-13
```




```
> confint(photo_m1)
                2.5 %   97.5 %
(Intercept) 4.708507 6.043222

> qnorm(c(0.025, 0.975),
+     mean = mu.hat, sd = mu.se)
[1] 4.746904 6.004826
```

lm: Confidence Intervals

```
> coef(summary(photo_m1))
            Estimate Std. Error  t value     Pr(>|t|)
(Intercept) 5.375865  0.3209046 16.75222 1.254999e-13
```

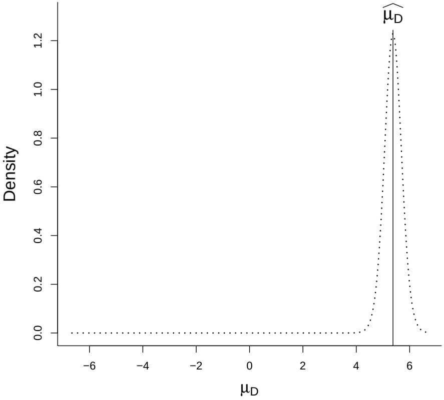



```
> confint(photo_m1)
                2.5 %   97.5 %
(Intercept) 4.708507 6.043222

> qnorm(c(0.025, 0.975),
+      mean = mu.hat, sd = mu.se)
[1] 4.746904 6.004826

> qt.scaled(c(0.025, 0.975),
+      mean = mu.hat, sd = mu.se,
+      df = 22 - 1)
[1] 4.708507 6.043222
```

`lm`: Hypothesis Testing

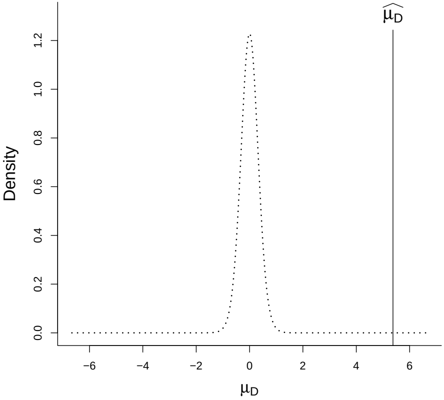
```
> coef(summary(photo_m1))
            Estimate Std. Error  t value     Pr(>|t|)
(Intercept) 5.375865  0.3209046 16.75222 1.254999e-13
```

```
> coef(summary(photo_m1))
             Estimate Std. Error  t value     Pr(>|t|)
(Intercept) 5.375865  0.3209046 16.75222 1.254999e-13
```



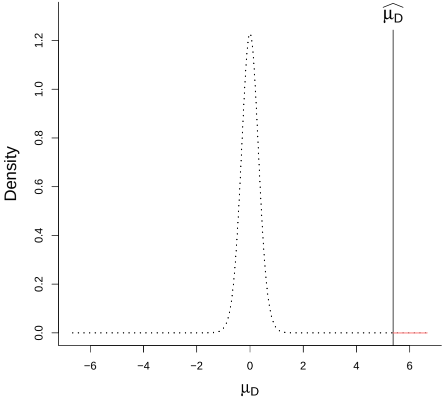
`lm`: Hypothesis Testing

```
> coef(summary(photo_m1))
             Estimate Std. Error  t value     Pr(>|t|)
(Intercept) 5.375865  0.3209046 16.75222 1.254999e-13
```



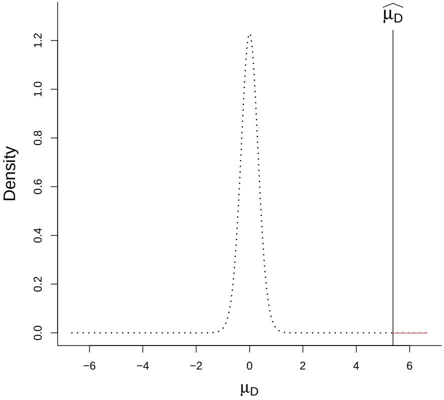
`lm`: Hypothesis Testing

```
> coef(summary(photo_m1))
             Estimate Std. Error  t value     Pr(>|t|)
(Intercept) 5.375865  0.3209046 16.75222 1.254999e-13
```



lm: Hypothesis Testing

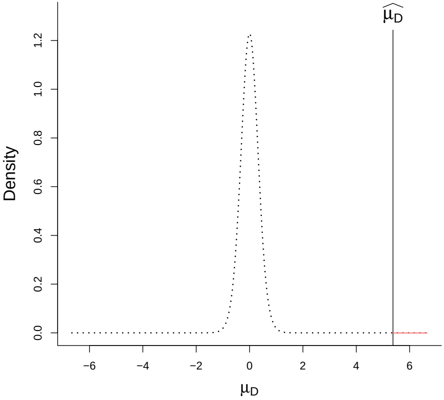
```
> coef(summary(photo_m1))
            Estimate Std. Error  t value     Pr(>|t|)
(Intercept) 5.375865  0.3209046 16.75222 1.254999e-13
```




```
> 1 - pt.scaled(mu.hat, mean = 0,
+       sd = mu.se, df = 22 - 1)
[1] 6.27276e-14
```

lm: Hypothesis Testing

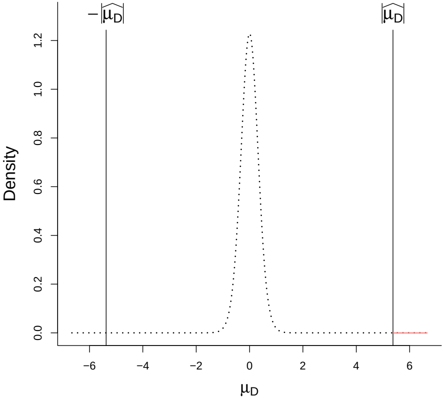
```
> coef(summary(photo_m1))
             Estimate Std. Error  t value     Pr(>|t|)
(Intercept) 5.375865  0.3209046 16.75222 1.254999e-13
```



One-tailed t-test

```
> 1 - pt.scaled(mu.hat, mean = 0,
+       sd = mu.se, df = 22 - 1)
[1] 6.27276e-14
```

```
> coef(summary(photo_m1))
            Estimate Std. Error  t value     Pr(>|t|)
(Intercept) 5.375865  0.3209046 16.75222 1.254999e-13
```

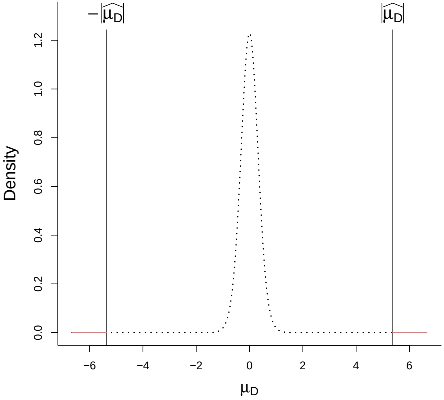


One-tailed t-test

```
> 1 - pt.scaled(mu.hat, mean = 0,
+     sd = mu.se, df = 22 - 1)
[1] 6.27276e-14
```

lm: Hypothesis Testing

```
> coef(summary(photo_m1))
              Estimate Std. Error  t value     Pr(>|t|)
(Intercept) 5.375865  0.3209046 16.75222 1.254999e-13
```

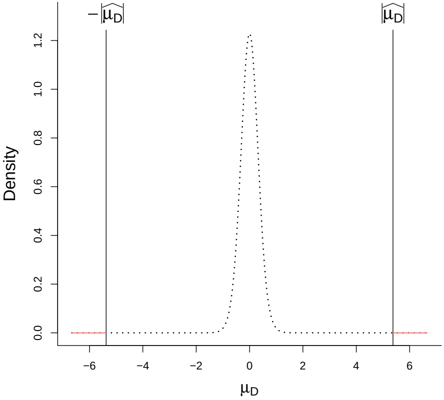


One-tailed t-test

```
> 1 - pt.scaled(mu.hat, mean = 0,
+       sd = mu.se, df = 22 - 1)
[1] 6.27276e-14
```

lm: Hypothesis Testing

```
> coef(summary(photo_m1))
             Estimate Std. Error  t value     Pr(>|t|)
(Intercept) 5.375865  0.3209046 16.75222 1.254999e-13
```

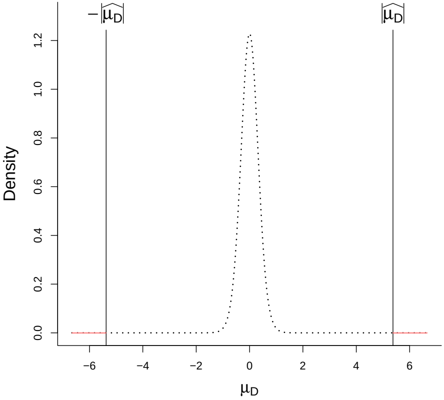


One-tailed t-test

```
> 1 - pt.scaled(mu.hat, mean = 0,
+       sd = mu.se, df = 22 - 1)
[1] 6.27276e-14

> pt.scaled(-mu.hat, mean = 0,
+       sd = mu.se, df = 22 - 1)
[1] 6.274995e-14
```

lm: Hypothesis Testing

```
> coef(summary(photo_m1))
            Estimate Std. Error  t value     Pr(>|t|)
(Intercept) 5.375865  0.3209046 16.75222 1.254999e-13
```



One-tailed t-test

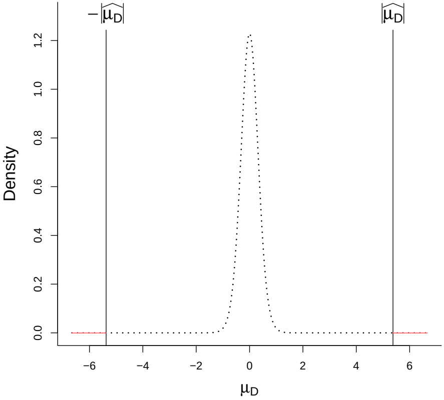
```
> 1 - pt.scaled(mu.hat, mean = 0,
+      sd = mu.se, df = 22 - 1)
[1] 6.27276e-14
```

```
> pt.scaled(-mu.hat, mean = 0,
+      sd = mu.se, df = 22 - 1)
[1] 6.274995e-14
```

Two-tailed t-test

lm: Hypothesis Testing

```
> coef(summary(photo_m1))
             Estimate Std. Error  t value     Pr(>|t|)
(Intercept) 5.375865  0.3209046 16.75222 1.254999e-13
```



One-tailed t-test

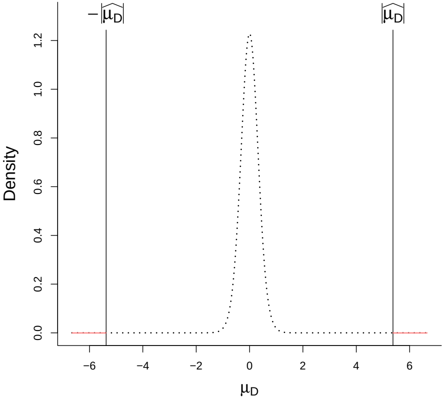
```
> 1 - pt.scaled(mu.hat, mean = 0,
+       sd = mu.se, df = 22 - 1)
[1] 6.27276e-14
```

```
> pt.scaled(-mu.hat, mean = 0,
+       sd = mu.se, df = 22 - 1)
[1] 6.274995e-14
```

Two-tailed t-test

```
> mu.t <- abs(mu.hat/mu.se)
```

```
> coef(summary(photo_m1))
            Estimate Std. Error  t value     Pr(>|t|)
(Intercept) 5.375865  0.3209046 16.75222 1.254999e-13
```



One-tailed t-test

```
> 1 - pt.scaled(mu.hat, mean = 0,
+       sd = mu.se, df = 22 - 1)
[1] 6.27276e-14
```

```
> pt.scaled(-mu.hat, mean = 0,
+       sd = mu.se, df = 22 - 1)
[1] 6.274995e-14
```

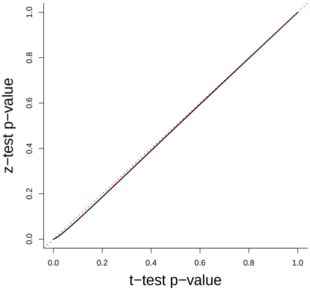
Two-tailed t-test

```
> mu.t <- abs(mu.hat/mu.se)
> 2 * (1 - pt(mu.t, df = 22 - 1))
[1] 1.254552e-13
```

lm: Hypothesis Testing

z-test

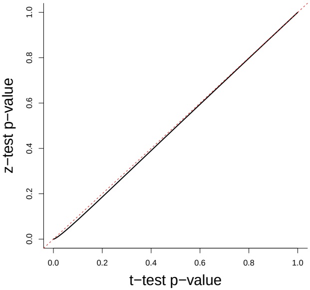
```
> 2 * (1 - pnorm(mu.t))
```



z-test

```
> 2 * (1 - pnorm(mu.t))
```

`lm`: Hypothesis Testing

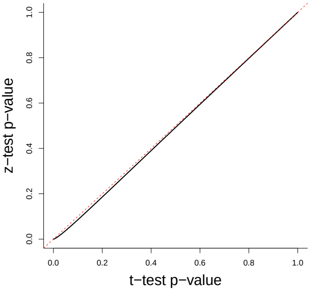


z-test

```
> 2 * (1 - pnorm(mu.t))
```

```
> logLik(photo_m1)
'log Lik.' -39.70094 (df=2)
```

lm: Hypothesis Testing

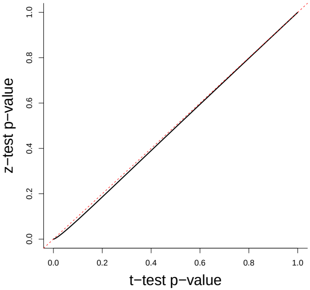


z-test

```
> 2 * (1 - pnorm(mu.t))
```

```
> logLik(photo_m1)
'log Lik.' -39.70094 (df=2)
> photo_m2 <- lm(y ~ -1, data = subset(photo_long,
+     type == "grumpy"))
```

lm: Hypothesis Testing



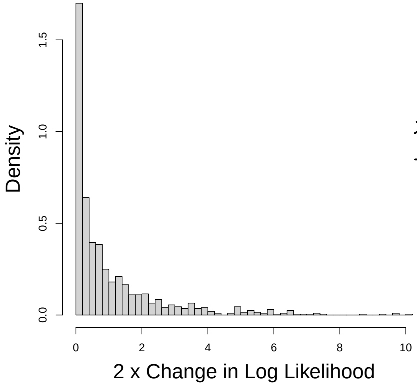
z-test

```
> 2 * (1 - pnorm(mu.t))
```

```
> logLik(photo_m1)
'log Lik.' -39.70094 (df=2)
> photo_m2 <- lm(y ~ -1, data = subset(photo_long,
+     type == "grumpy"))
> logLik(photo_m2)
'log Lik.' -69.01266 (df=1)
```

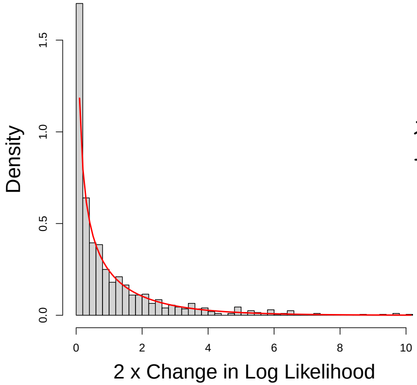
Likelihood Ratio Test

```
> LR2 <- 2 * (logLik(photo_m1) -
+     logLik(photo_m2))
```



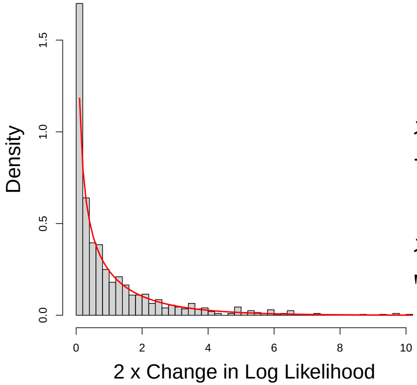
Likelihood Ratio Test

```
> LR2 <- 2 * (logLik(photo_m1) -
+       logLik(photo_m2))
```



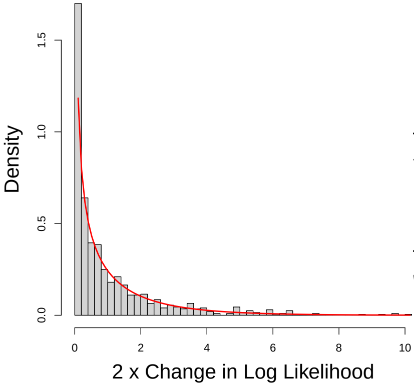
Likelihood Ratio Test

```
> LR2 <- 2 * (logLik(photo_m1) -
+      logLik(photo_m2))
```



Likelihood Ratio Test

```
> LR2 <- 2 * (logLik(photo_m1) -
+       logLik(photo_m2))

> 1 - pchisq(LR2, df = 1)
'log Lik.' 1.909584e-14 (df=2)
```



Likelihood Ratio Test

```
> LR2 <- 2 * (logLik(photo_m1) -
+     logLik(photo_m2))
```

```
> 1 - pchisq(LR2, df = 1)
'log Lik.' 1.909584e-14 (df=2)
```

```
> anova(photo_m1, photo_m2, test = "LRT")
  Res.Df    RSS Df Sum of Sq  Pr(>Chi)
1     22 683.37
2     21  47.58  1     635.8 < 2.2e-16 ***
---
```

Summary

- Distributions
  - Data Distribution - probability of data.
  - Sampling Distribution - probability of estimates.
  - Posterior Distribution - probability (epistemic) of parameter values.
- Distribution Functions
  - Mass/Density - (proportional to the) probability that $X = x$.
  - Cumulative Density/Mass - probability that $X \leq x$.
  - Quantile - Opposite of Cumulative: return probability given x.
- Inference
  - Maximum Likelihood (ML) - choose parameters that maximise the probability of the data given the model.
  - If the data are Gaussian we have the linear model.
    - Sampling distribution of location parameters are t-distributed (close to Gaussian in many cases).
    - Hypothesis testing with t-test (close to z-test in many cases).
  - Likelihood ratio test for general hypothesis testing under ML.