Quantitative Genetics, Prediction and Selection Theory

SELECTION INDICES

BLUP and REML

BAYESIAN INFERENCE

GENOMIC SELECTION

BLUP and REML

1. The linear model

The fixed effects model

Interlude: Fixed effects and random effects

The mixed model

2. BLUP

The mixed model equations

Genetic trends

Other models

BLUP Properties

3. Estimation of genetic parameters

The maximum likelihood method

REML

Genetic parameters estimation

The Linear Model

In selection indices, the data are **centred**

THE FIXED EFFECTS MODEL

$$y = \mu + E + P + \beta x + e = \boldsymbol{Xb} + \boldsymbol{e}$$

$$E(\boldsymbol{e}') = E[e_1, e_2, \ldots, e_n] = [0, 0, \ldots, 0] \qquad \text{var}(\mathbf{e}) = \begin{bmatrix} \sigma^2_{e_1} & 0 & \cdots & 0 \\ & \sigma^2_{e_2} & \cdots & 0 \\ & & \ddots & \vdots \\ & & & \sigma^2_{e_n} \end{bmatrix} = \mathbf{I}\sigma^2_e$$

$$\boxed{\hat{\mathbf{b}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}}$$

The Linear Model

THE FIXED EFFECTS MODEL

$$12 = \mu + E_1 + P_1 + \beta \cdot 20 + e_1$$
$$7 = \mu + E_2 + P_2 + \beta \cdot 9 + e_2$$
$$9 = \mu + E_1 + P_1 + \beta \cdot 15 + e_3$$
$$11 = \mu + E_2 + P_2 + \beta \cdot 16 + e_4$$
$$8 = \mu + E_2 + P_3 + \beta \cdot 14 + e_5$$

… … … … … … … … … …

… … … … … … … … … …

$$\begin{bmatrix} 12 \\ 7 \\ 9 \\ 11 \\ 8 \\ \vdots \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 & 0 & 20 \\ 1 & 0 & 1 & 0 & 1 & 0 & 9 \\ 1 & 1 & 0 & 1 & 0 & 0 & 15 \\ 1 & 0 & 1 & 0 & 1 & 0 & 16 \\ 1 & 0 & 1 & 0 & 0 & 1 & 14 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{bmatrix} \cdot \begin{bmatrix} \mu \\ E_1 \\ E_2 \\ P_1 \\ P_2 \\ P_3 \\ \beta \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ \vdots \end{bmatrix}$$

$$\mathbf{y} = \mathbf{X} \quad \mathbf{b} + \mathbf{e}$$

The Linear Model

THE FIXED EFFECTS MODEL

$$\begin{bmatrix} 12 \\ 7 \\ 9 \\ 11 \\ 8 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ E_1 \\ E_2 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \end{bmatrix}$$

$$\mathbf{X'X} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 2 & 3 \\ 2 & 2 & 0 \\ 3 & 0 & 3 \end{bmatrix}$$

$$\boxed{E_1 + E_2 = 0}$$

$$\begin{bmatrix} 12 \\ 7 \\ 9 \\ 11 \\ 8 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} \mu \\ E_1 \\ E_2 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \end{bmatrix}$$

$$\mathbf{X'X} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 4 \end{bmatrix}$$

The Linear Model

THE FIXED EFFECTS MODEL

$$\begin{bmatrix} 12 \\ 7 \\ 9 \\ 11 \\ 8 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ E_1 \\ E_2 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \end{bmatrix}$$

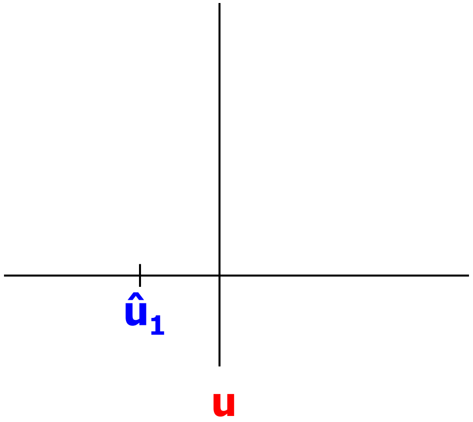
$$\boldsymbol{X'X} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 2 & 3 \\ 2 & 2 & 0 \\ 3 & 0 & 3 \end{bmatrix}$$

$$\begin{bmatrix} 12 \\ 7 \\ 9 \\ 11 \\ 8 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \mu + E_1 \\ \mu + E_2 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \end{bmatrix} \quad \rightarrow \quad \boldsymbol{X'X} = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$$

Interlude:

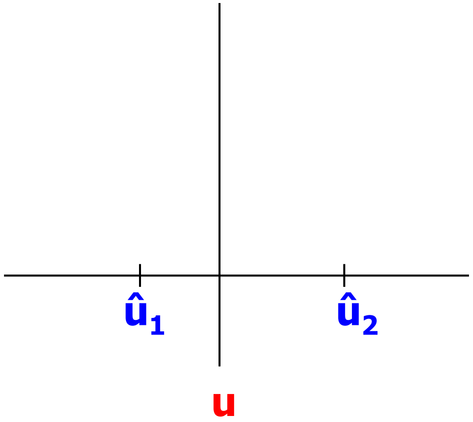
Fixed effects and random effects

Fixed or random?



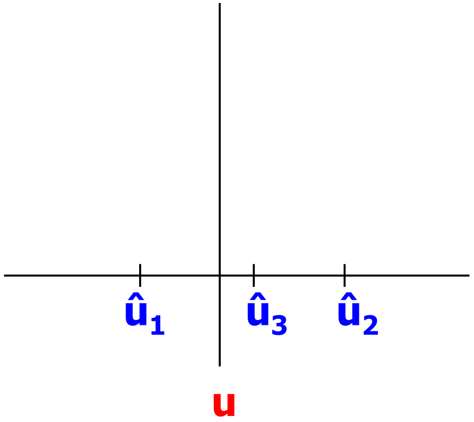
FIXED

Fixed or random?



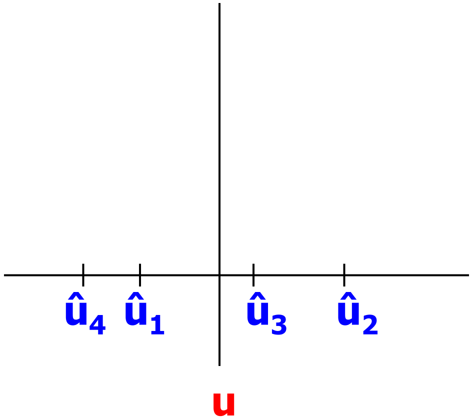
FIXED

Fixed or random?



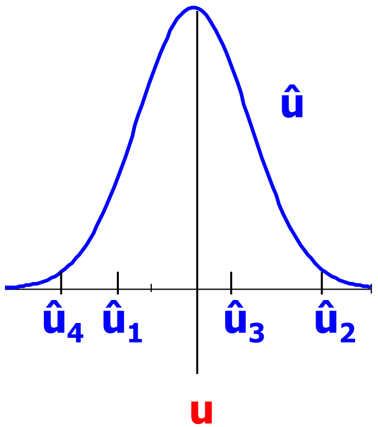
FIXED

Fixed or random?



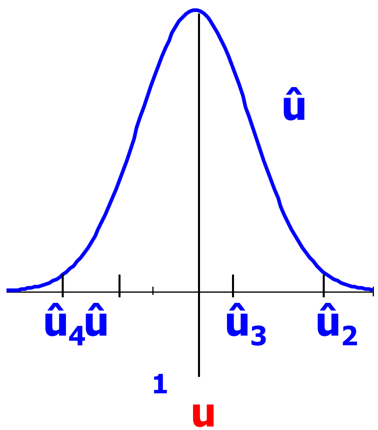
FIXED

Fixed or random?



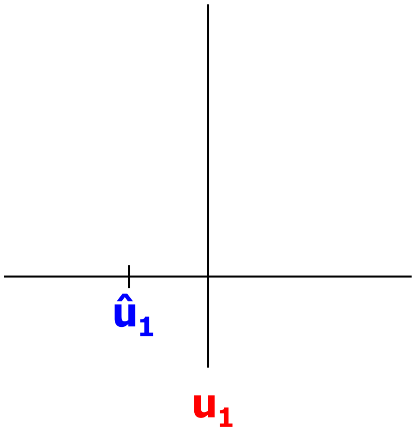
FIXED

Fixed or random?

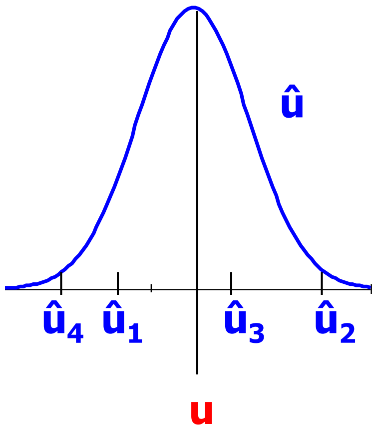


FIXED

RANDOM

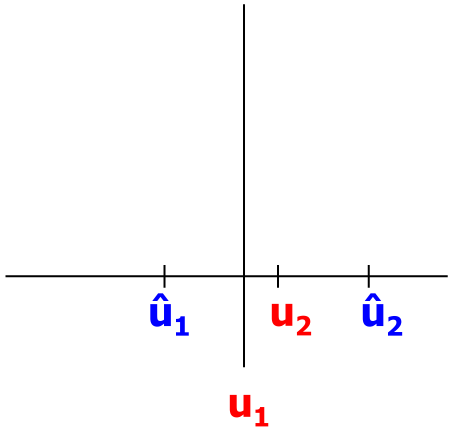


Fixed or random?

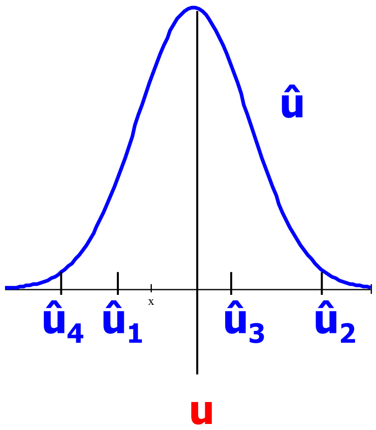


FIXED

RANDOM

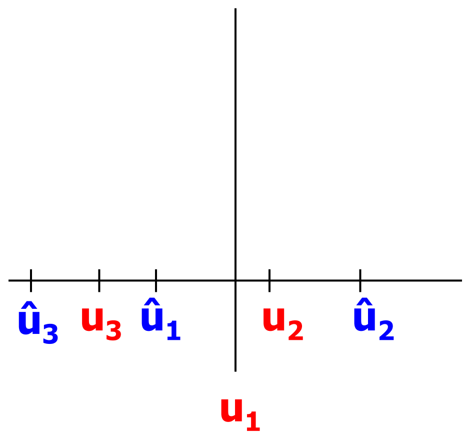


Fixed or random?

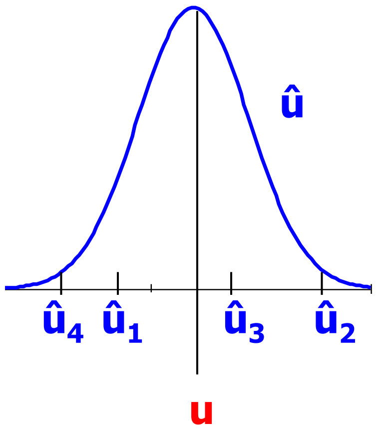


FIXED

RANDOM

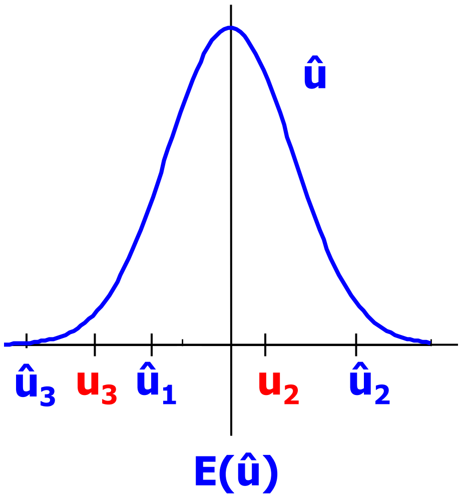


Fixed or random?

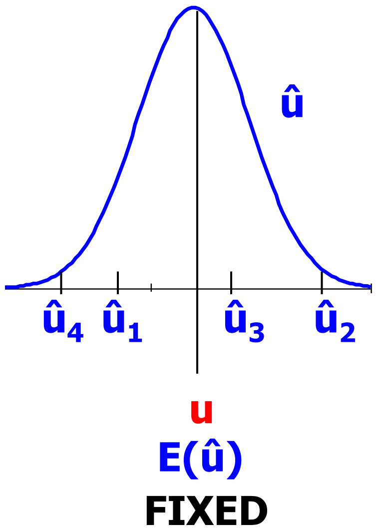


FIXED

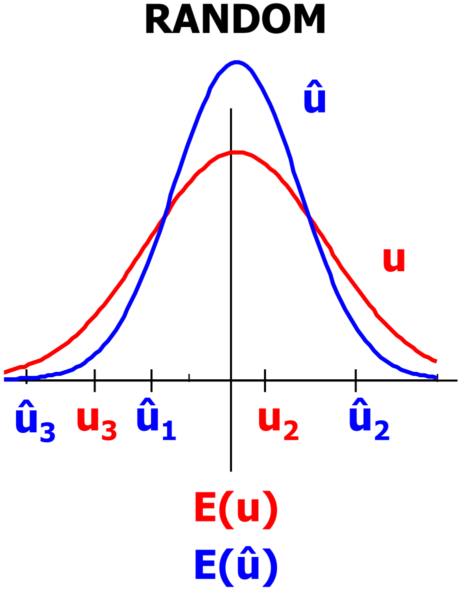
RANDOM



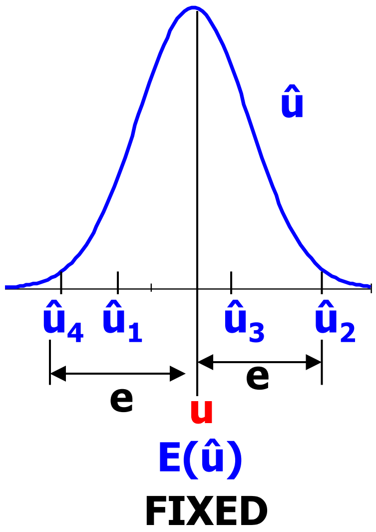
Fixed or random?



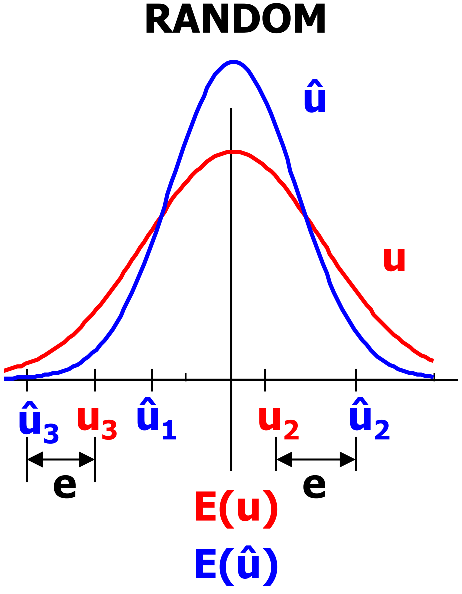
To be "unbiased" is less attractive



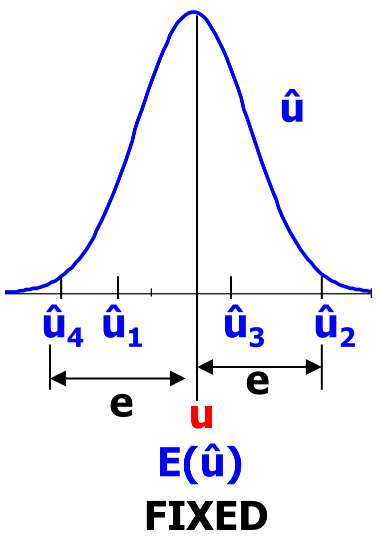
Fixed or random?

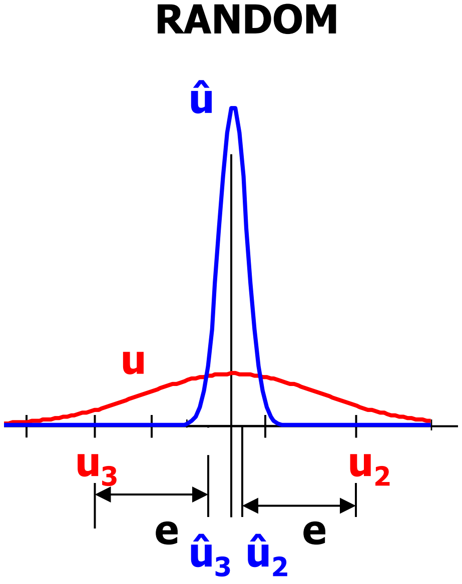


To be "unbiased"
is less attractive



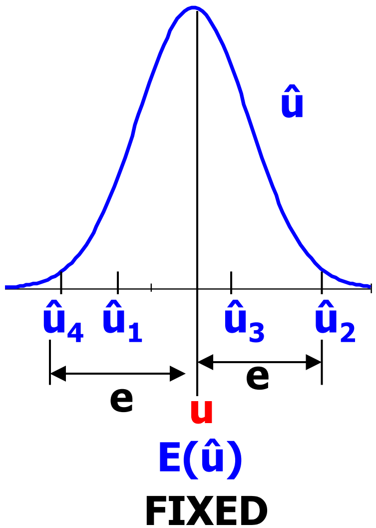
Fixed or random?

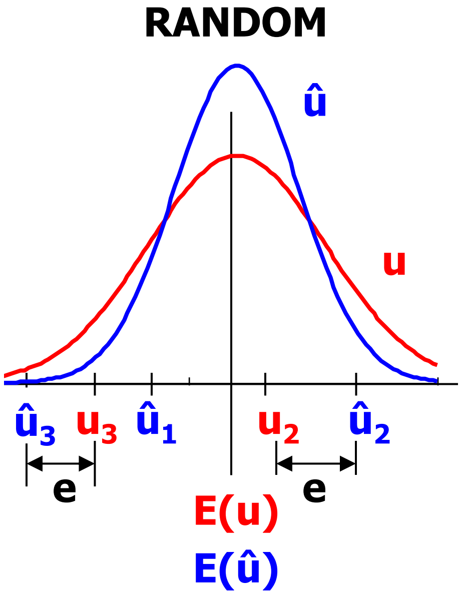




Fixed or random?

When **n** increases,
FIXED: var(û) decreases
RANDOM: var(û) increases





Fixed or random?

Problem: Evaluate a bull with n daughters from different cows

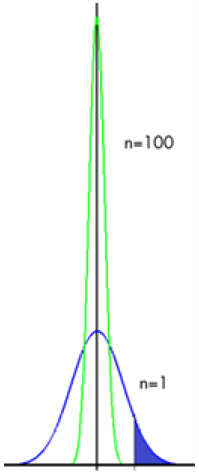
FIXED

$$\hat{A}_{FIXED} = \bar{D}$$

RANDOM

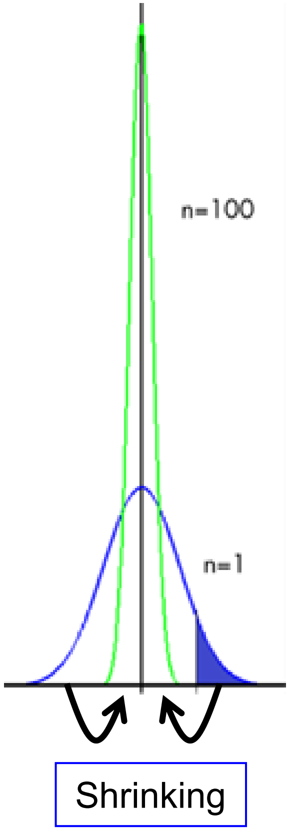
$$\hat{A}_{RANDOM} = b \cdot \bar{D} = \frac{cov(A,\bar{D})}{\sigma^2_{\bar{D}}} \cdot \bar{D} = \frac{\frac{1}{2}\sigma^2_A}{\sigma^2_{\bar{D}}} \cdot \bar{D}$$

$$\hat{A}_{RANDOM} \propto \frac{\bar{D}}{\sigma^2_{\bar{D}}}$$



Fixed or random?

$$\hat{A}_{RANDOM} \propto \frac{\bar{D}}{\sigma^2_{\bar{D}}}$$



Fixed or random?

- Breed effects: Fixed or random?
- Herd effects: Fixed or radom?
- Sire effects: Fixed or random?
- Treatment (panelist): Fixed or random?
- Treatment (diet): Fixed or random?

Fixed or random?

- FIXED: Large n
  Few levels
  Leads to a good correction

- RANDOM: Small n
  Many levels
  Leads to poor correction
  Leads to a better residuals structure

End of the Interlude

The Linear Model

THE MIXED MODEL

$$y = \mu + E + P + a + e$$

$$y = \mu + E + P + \beta x + a + p + e$$

y: data from a single lactation (e.g., kg of milk)
E: season
P: calving (lactation number)
x: age of the cow
a: additive genetic effect
p: permanent environmental effect across lactations
e: residual

The Linear Model

THE MIXED MODEL

$$y = \mu + E + P + a + e$$

$$y = \mu + E + P + \beta x + a + p + e$$

$$\boldsymbol{a} \sim N\left(\boldsymbol{0}, \boldsymbol{A}\sigma_A^2\right) \quad \boldsymbol{p} \sim N\left(\boldsymbol{0}, \boldsymbol{I}\sigma_p^2\right) \quad \boldsymbol{e} \sim N\left(\boldsymbol{0}, \boldsymbol{I}\sigma_e^2\right)$$

$$\mathrm{var}\left(\mathbf{a}\right) = \begin{bmatrix} \sigma_A^2 & \mathrm{cov}\left(a_1, a_2\right) & \cdots & \mathrm{cov}\left(a_1, a_m\right) \\ & \sigma_A^2 & \cdots & \mathrm{cov}\left(a_2, a_m\right) \\ & & \ddots & \vdots \\ & & & \sigma_A^2 \end{bmatrix} = \begin{bmatrix} 1 & 2r_{12} & \cdots & 2r_{1m} \\ & 1 & \cdots & 2r_{2m} \\ & & \ddots & \vdots \\ & & & 1 \end{bmatrix} \sigma_A^2 = \mathbf{A}\sigma_A^2$$

$$\mathrm{cov}\left(\boldsymbol{a}, \boldsymbol{p}\right) = \boldsymbol{0};\ \mathrm{cov}\left(\boldsymbol{a}, \boldsymbol{e}\right) = \boldsymbol{0};\ \mathrm{cov}\left(\boldsymbol{p}, \boldsymbol{e}\right) = \boldsymbol{0}$$

The Linear Model

THE MIXED MODEL

$$
\begin{aligned}
12 &= \mu + E_1 + P_1 + \beta \cdot 20 + a_1 + p_1 + e_1 \\
7 &= \mu + E_2 + P_2 + \beta \cdot \ 9 + a_1 + p_1 + e_2 \\
9 &= \mu + E_1 + P_1 + \beta \cdot 15 + a_2 + p_2 + e_3 \\
11 &= \mu + E_2 + P_2 + \beta \cdot 16 + a_2 + p_2 + e_4 \\
8 &= \mu + E_2 + P_3 + \beta \cdot 14 + a_2 + p_2 + e_5 \\
&\cdots \quad \cdots \quad \cdots \quad \cdots
\end{aligned}
$$

$$
\begin{bmatrix} 12 \\ 7 \\ 9 \\ 11 \\ 8 \\ \vdots \end{bmatrix}
=
\begin{bmatrix}
1 & 1 & 0 & 1 & 0 & 0 & 20 \\
1 & 0 & 1 & 0 & 1 & 0 & 9 \\
1 & 1 & 0 & 1 & 0 & 0 & 15 \\
1 & 0 & 1 & 0 & 1 & 0 & 16 \\
1 & 0 & 1 & 0 & 0 & 1 & 14 \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots
\end{bmatrix}
\cdot
\begin{bmatrix} \mu \\ E_1 \\ E_2 \\ P_1 \\ P_2 \\ P_3 \\ \beta \end{bmatrix}
+
\begin{bmatrix}
1 & 0 & \cdots & 0 \\
1 & 0 & \cdots & 0 \\
0 & 1 & \cdots & 0 \\
0 & 1 & \cdots & 0 \\
0 & 1 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots
\end{bmatrix}
\begin{bmatrix} a_1 \\ a_2 \\ \vdots \end{bmatrix}
+
\begin{bmatrix}
1 & 0 & \cdots & 0 \\
1 & 0 & \cdots & 0 \\
0 & 1 & \cdots & 0 \\
0 & 1 & \cdots & 0 \\
0 & 1 & \cdots & 0 \\
\vdots & \vdots & \vdots & \vdots
\end{bmatrix}
\begin{bmatrix} p_1 \\ p_2 \\ \vdots \end{bmatrix}
+
\begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \\ \vdots \end{bmatrix}
$$

$$
\mathbf{y} = \mathbf{X} \quad \mathbf{b} + \mathbf{Z} \quad \mathbf{a} + \mathbf{W} \quad \mathbf{p} + \mathbf{e}
$$

BLUP

$$\boldsymbol{y} = \boldsymbol{X}\boldsymbol{b} + \boldsymbol{Z}\boldsymbol{a} + \boldsymbol{e}$$

$$\boldsymbol{y} - \boldsymbol{X}\boldsymbol{b} = \boldsymbol{Z}\boldsymbol{a} + \boldsymbol{e}$$

$$\tilde{\boldsymbol{a}} = \boldsymbol{C}'\boldsymbol{V}^{-1}\left(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{b}\right)$$

$$\hat{\boldsymbol{b}} = \left(\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{X}\right)^{-1}\boldsymbol{X}'\boldsymbol{V}^{-1}\boldsymbol{y}$$

$$\hat{\boldsymbol{a}} = \boldsymbol{C}'\boldsymbol{V}^{-1}\left(\boldsymbol{y} - \boldsymbol{X}\hat{\boldsymbol{b}}\right)$$

$$\begin{bmatrix} \boldsymbol{X}'\boldsymbol{X} & \boldsymbol{X}'\boldsymbol{Z} \\ \boldsymbol{Z}'\boldsymbol{X} & \boldsymbol{Z}'\boldsymbol{Z} + \boldsymbol{A}^{-1}\dfrac{\sigma_e^2}{\sigma_A^2} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{b}} \\ \hat{\boldsymbol{a}} \end{bmatrix} = \begin{bmatrix} \boldsymbol{X}'\boldsymbol{y} \\ \boldsymbol{Z}'\boldsymbol{y} \end{bmatrix}$$

BLUP

EXAMPLE

$$
\begin{aligned}
12 &= \mu + E_1 + a_1 + e_1 \\
7 &= \mu + E_2 + a_1 + e_2 \\
9 &= \mu + E_1 + a_2 + e_3 \\
11 &= \mu + E_2 + a_2 + e_4 \\
8 &= \mu + E_2 + a_2 + e_5
\end{aligned}
$$

$$
\begin{bmatrix} 12 \\ 7 \\ 9 \\ 11 \\ 8 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} \mu + E_1 \\ \mu + E_2 \end{bmatrix} + \begin{bmatrix} 1 & 0 & \boxed{0} \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ e_3 \\ e_4 \\ e_5 \end{bmatrix}
$$

$$
\boldsymbol{y} = \boldsymbol{X} \quad \boldsymbol{b} + \boldsymbol{Z} \quad \boldsymbol{a} + \boldsymbol{e}
$$

Females a_1 and a_2 are half-sisters. Male a_3 is a full brother of a_1 and a half-brother of a_2

$$
\boldsymbol{A} = \begin{bmatrix} 1 & 0.25 & 0.5 \\ & 1 & 0.25 \\ & & 1 \end{bmatrix} \quad \rightarrow \quad \boldsymbol{A}^{-1} = \begin{bmatrix} 1.36 & -0.18 & -0.64 \\ & 1.09 & -0.18 \\ & & 1.36 \end{bmatrix}
$$

$$
\frac{\sigma_e^2}{\sigma_A^2} = \frac{\sigma_P^2 - \sigma_A^2}{\sigma_A^2} = \frac{\sigma_P^2 - \sigma_P^2 h^2}{\sigma_P^2 h^2} = \frac{1 - h^2}{h^2} = 9 \quad \rightarrow \quad \boldsymbol{A}^{-1} \cdot 9 = \begin{bmatrix} 12 & -1.6 & -5.7 \\ & 9.8 & -1.6 \\ & & 12 \end{bmatrix}
$$

BLUP

EXAMPLE

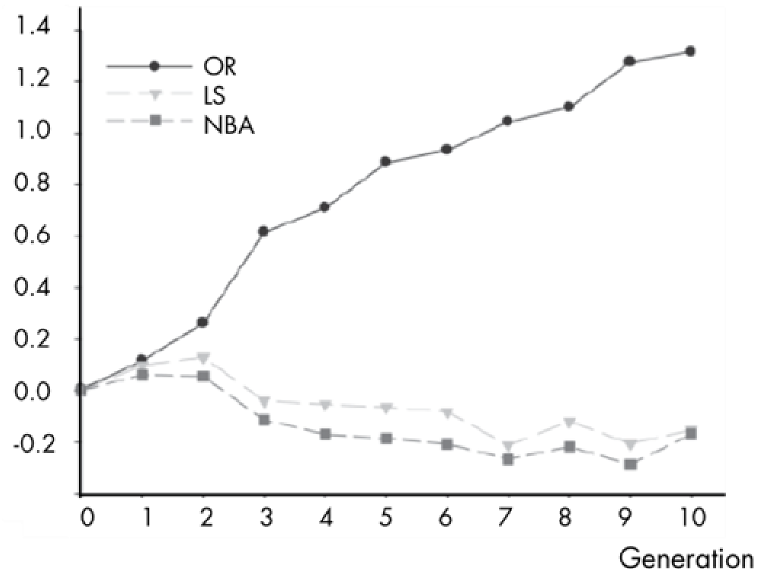
$$\begin{bmatrix} X'X & X'Z \\ Z'X & Z'Z + A^{-1}\dfrac{\sigma_e^2}{\sigma_A^2} \end{bmatrix} \begin{bmatrix} \hat{b} \\ \hat{a} \end{bmatrix} = \begin{bmatrix} X'y \\ Z'y \end{bmatrix}$$

$$\begin{bmatrix} 2 & 0 & 1 & 1 & 0 & & & \\ 0 & 3 & 1 & 2 & 0 & & & \\ & & & & & & & \\ 1 & 1 & 2 & 0 & 0 & 12 & -1.6 & -5.7 \\ 1 & 2 & 0 & 3 & 0 \quad + & -1.6 & 9.8 & 1.6 \\ 0 & 0 & 0 & 0 & 0 & -5.7 & -1.6 & 12 \end{bmatrix} \begin{bmatrix} \mu + E_1 \\ \mu + E_2 \\ \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 21 \\ 26 \\ \\ 19 \\ 28 \\ 0 \end{bmatrix}$$

$$\begin{bmatrix} 2 & 0 & 1 & 1 & 0 \\ 0 & 3 & 1 & 2 & 0 \\ 1 & 1 & 14 & -1.6 & -5.7 \\ 1 & 2 & -1.6 & 12.8 & -1.6 \\ 0 & 0 & -5.7 & -1.6 & 12 \end{bmatrix} \begin{bmatrix} \mu + E_1 \\ \mu + E_2 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 21 \\ 26 \\ 19 \\ 28 \\ 0 \end{bmatrix} \quad \rightarrow \quad \begin{bmatrix} \mu + E_1 \\ \mu + E_2 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 10.50 \\ 8.66 \\ -0.012 \\ 0.012 \\ -0.004 \end{bmatrix}$$

BLUP

Genetic trends



- The complete relationship matrix A from generation zero should be used
- All the data used in selection should be analyzed together
- The true genetic parameters *in the base generation* must be used

BLUP

Other models

1. Multi-trait BLUP $a_E = w_1 a_1 + \cdots + w_m a_m$
2. Maternal Effects Model $y = \mu + F + a + m_a + m_p + e$
3. Sire Model $y = \mu + F + s + e$
4. Model with dominance $y = \mu + F + a + d + e$
5. Model with genetic groups $y = \mu + F + G + a + e$
6. Reduced animal model

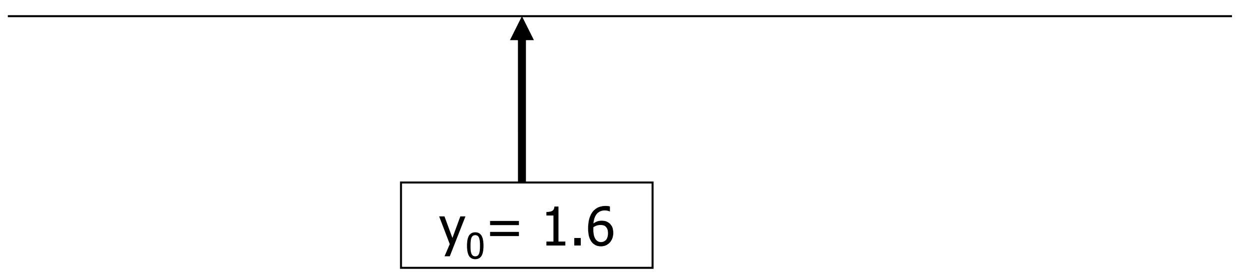
BLUP

Properties of BLUP

The properties of selection indices also apply to BLUP

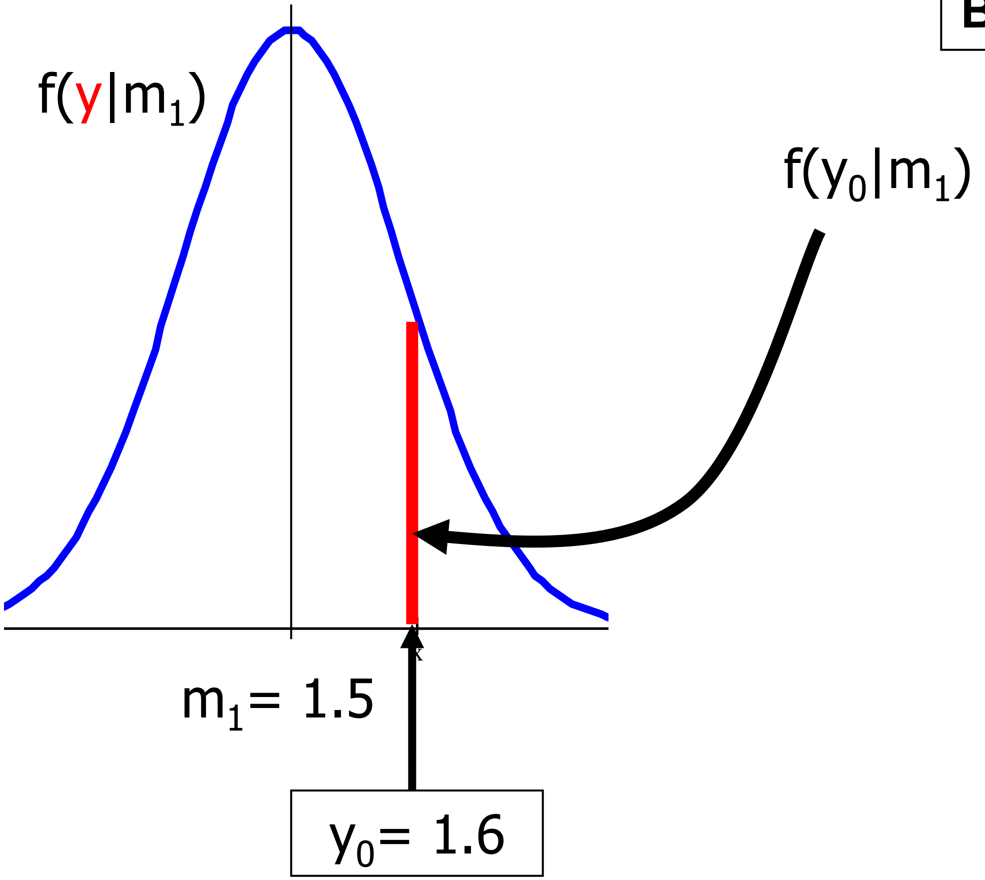
- BLUP maximizes the response to selection
  If the genetic parameters used to calculate it are the true ones
- BLUP maximizes the correlation between the predicted and the true value
- BLUP has the minimum risk among the unbiased estimators
- The prediction errors increase with the number of generations
  this is due to the increase in relationships
- BLUP predictions are not affected by the reduction in variance due to selection
  since they only use the variances of the base population
- Data should not necessarily be Normally distributed to apply BLUP

Likelihood

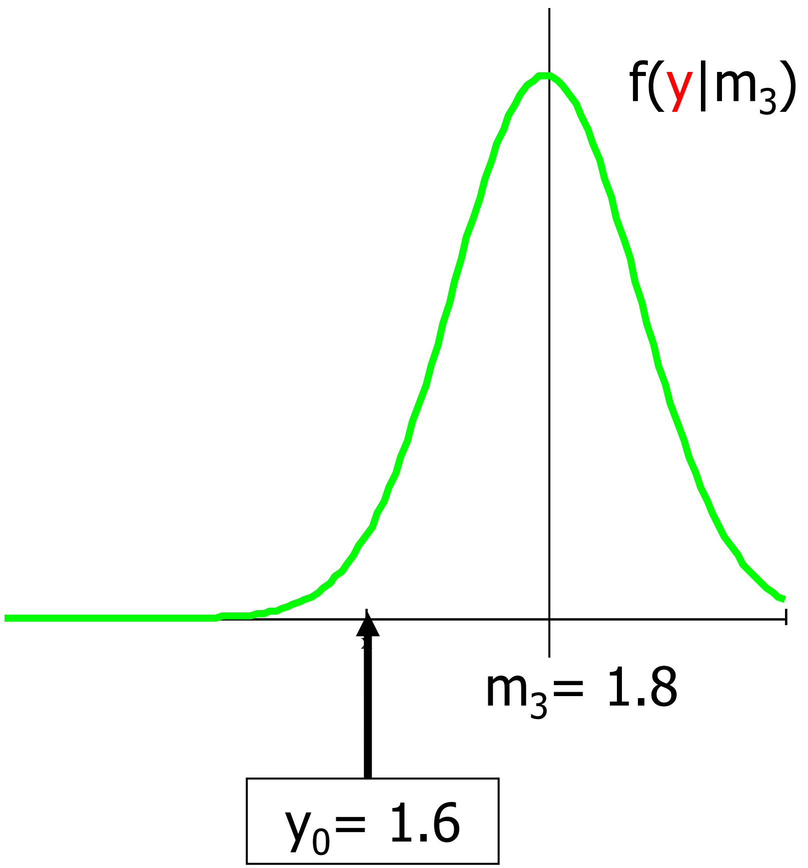


Likelihood

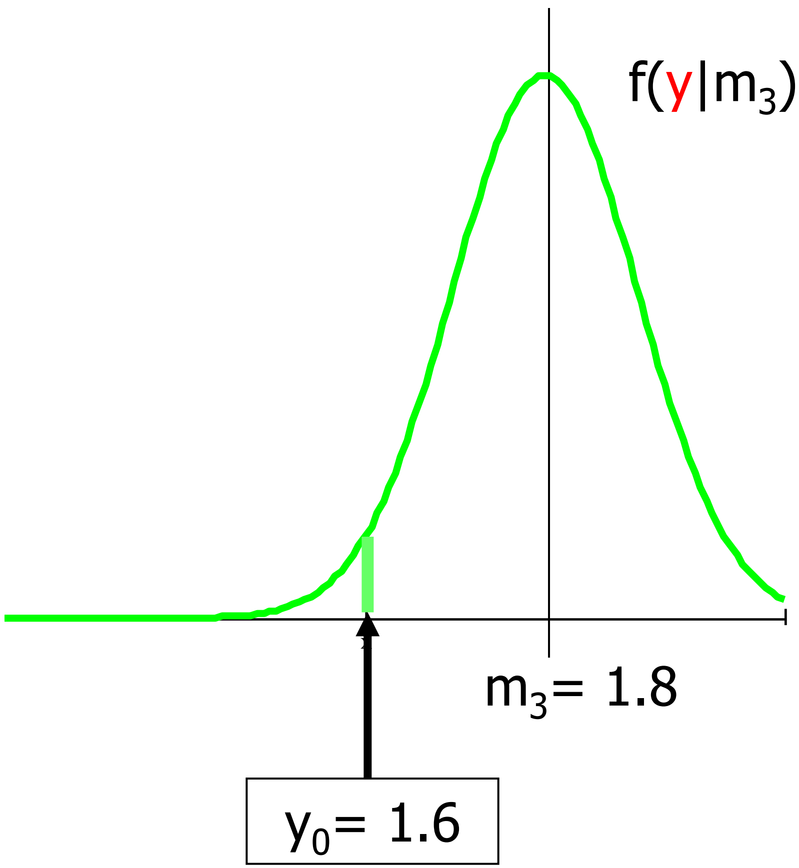
NOTATION
Red: Variable
Black: Constant



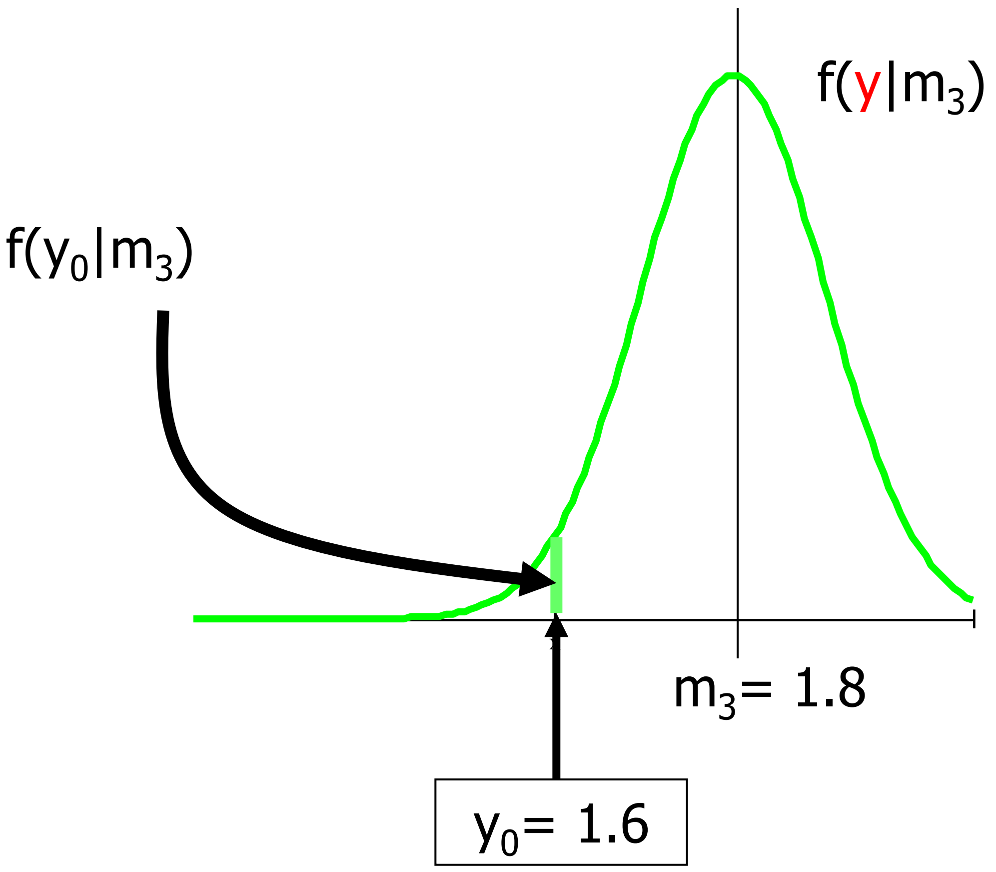
Likelihood



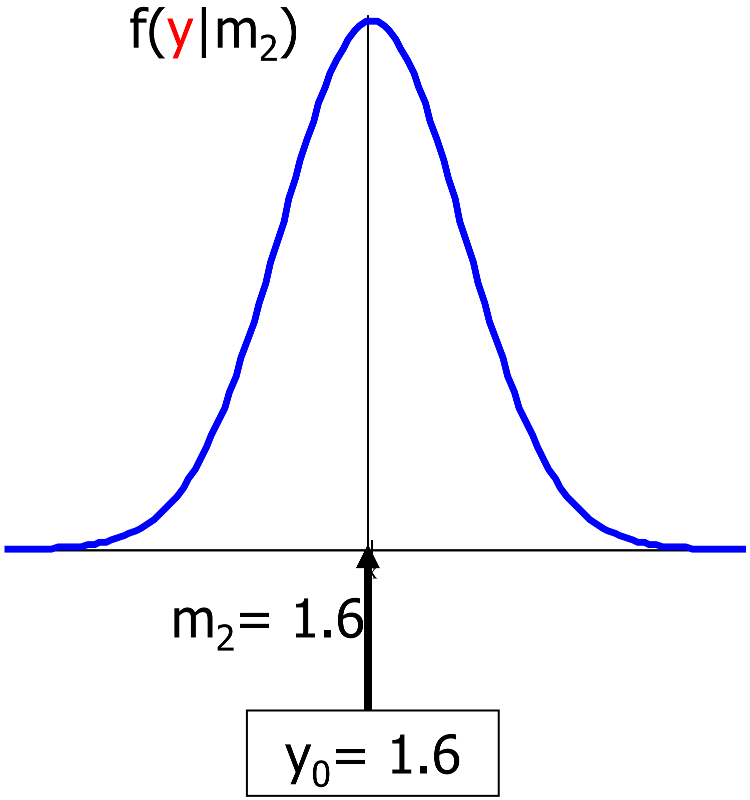
Likelihood



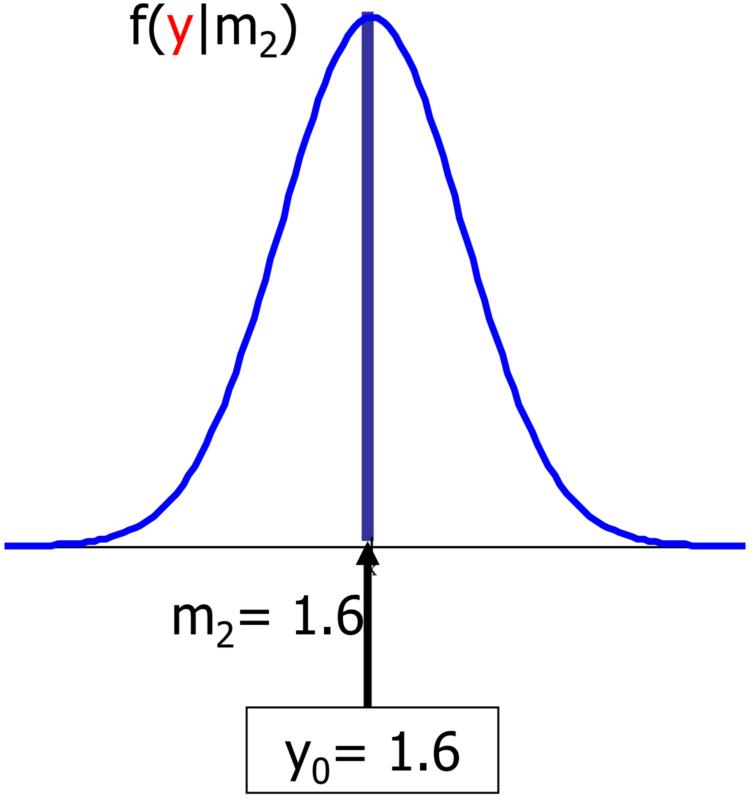
Likelihood



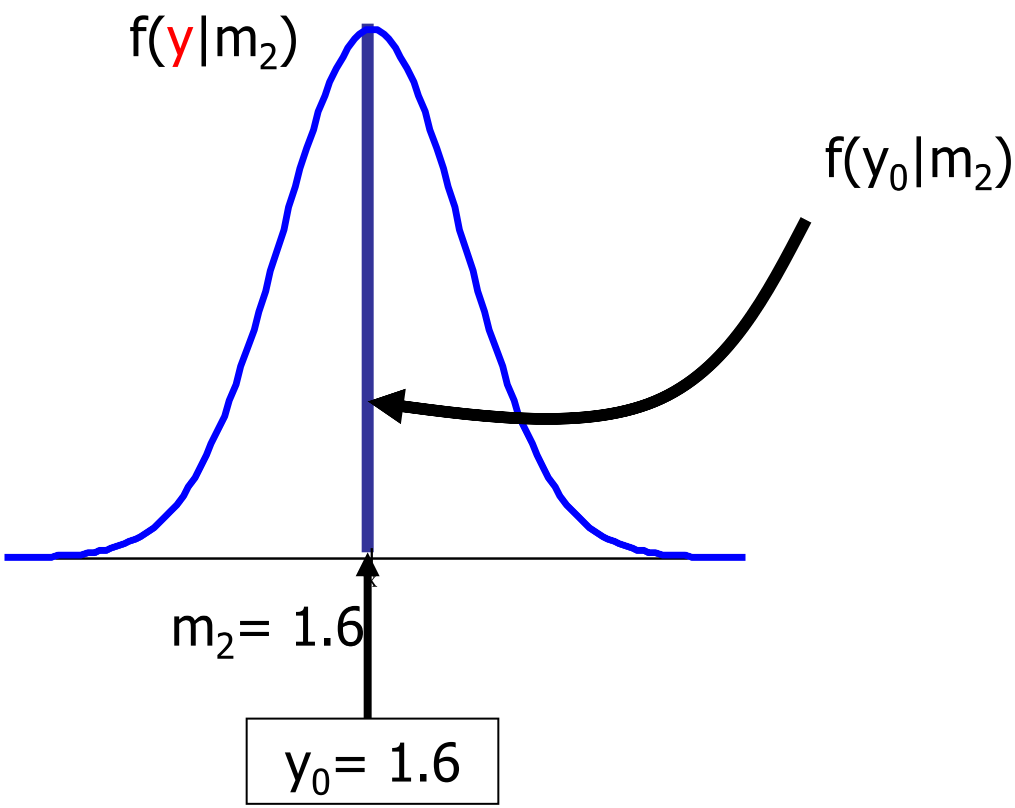
Likelihood



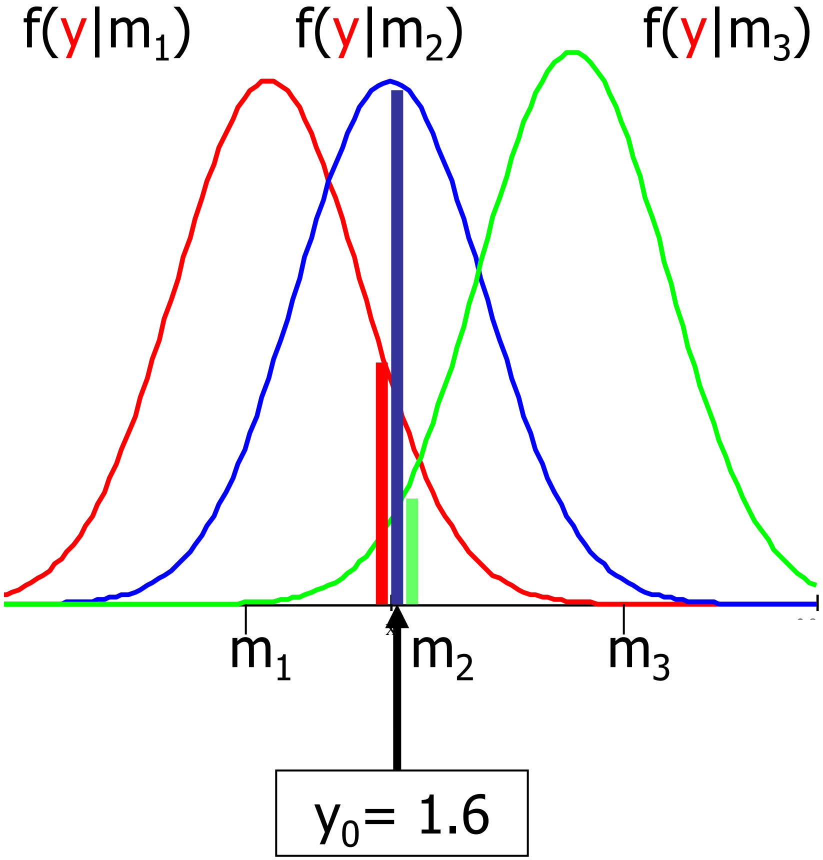
Likelihood



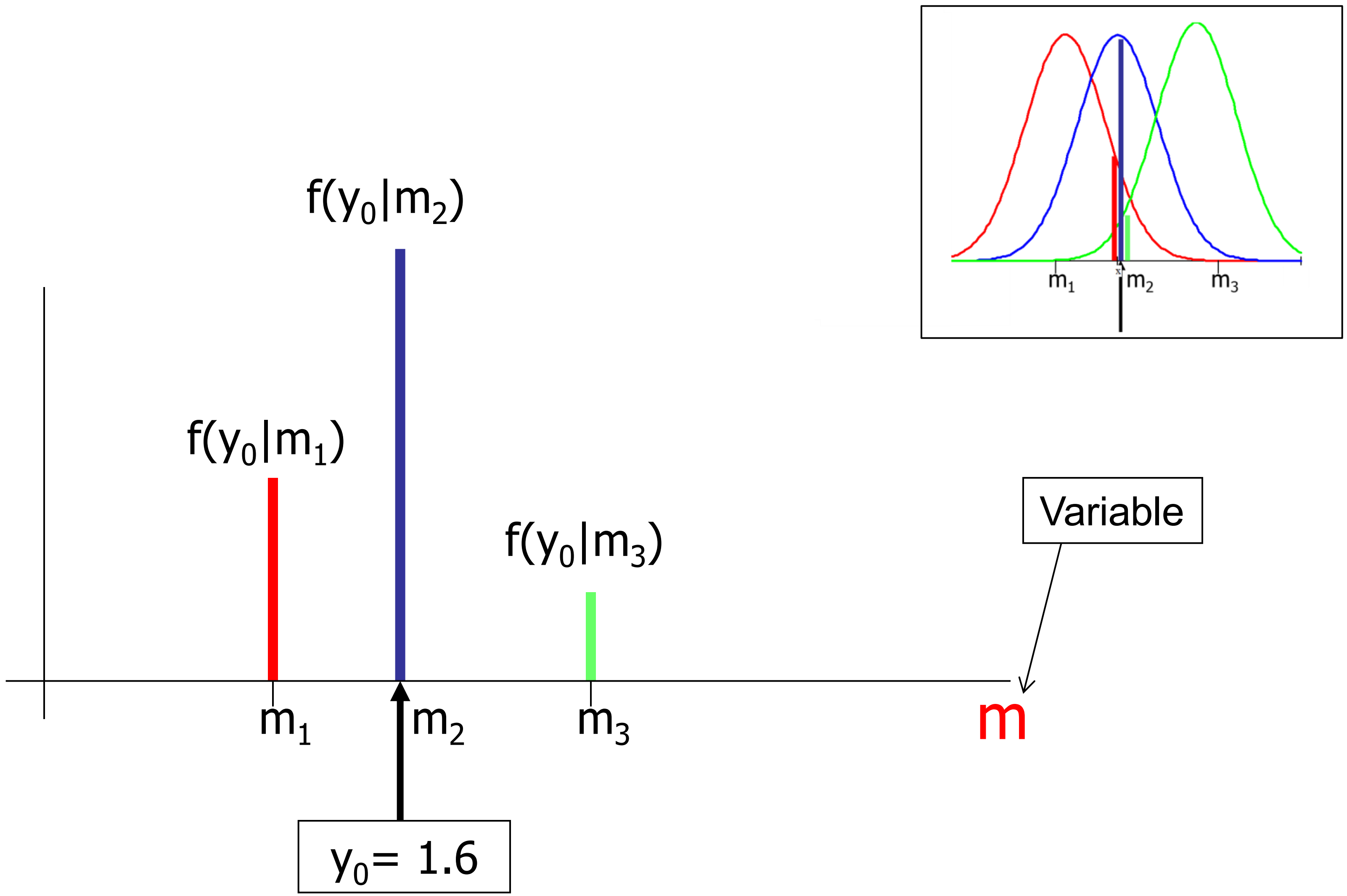
Likelihood



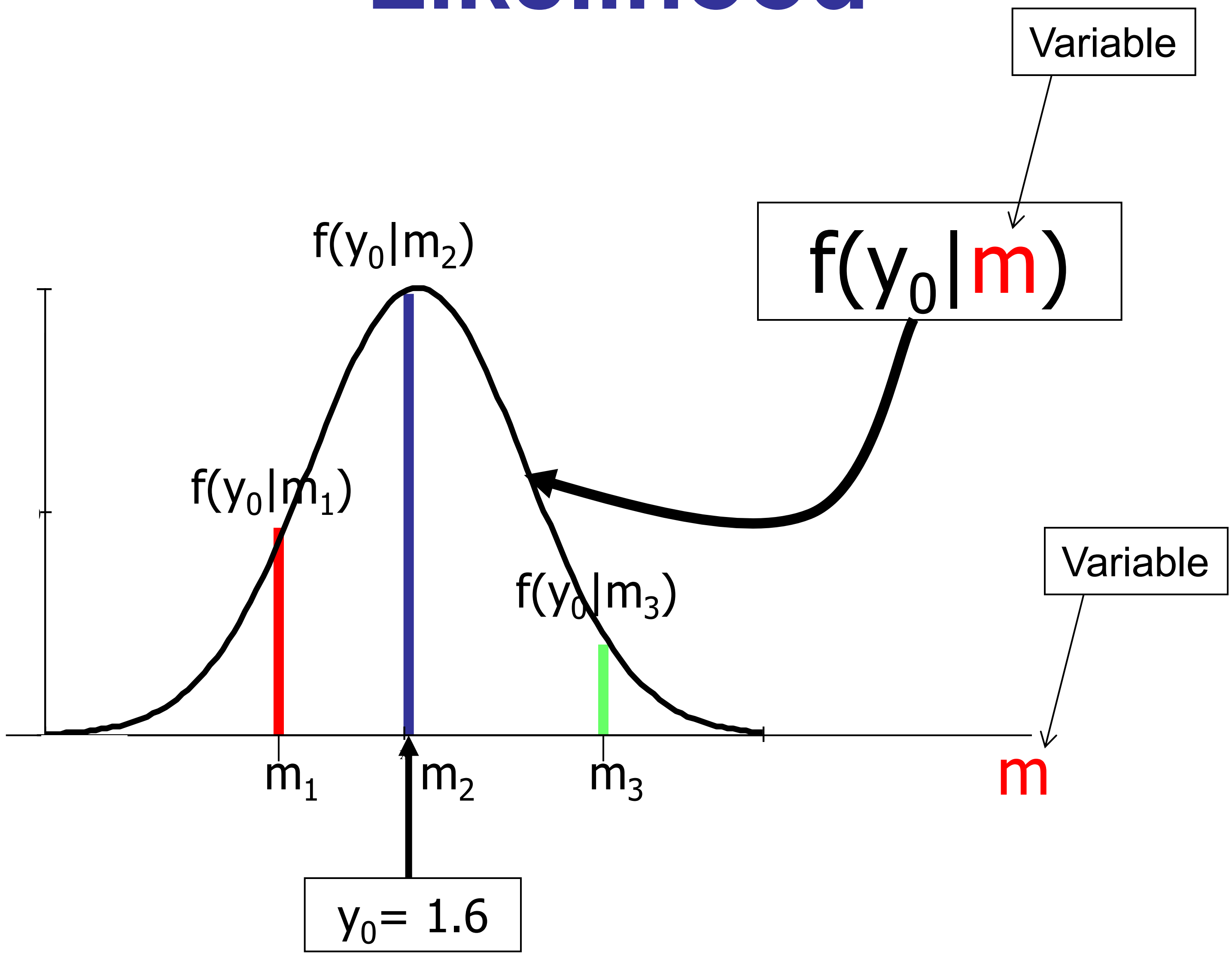
Likelihood



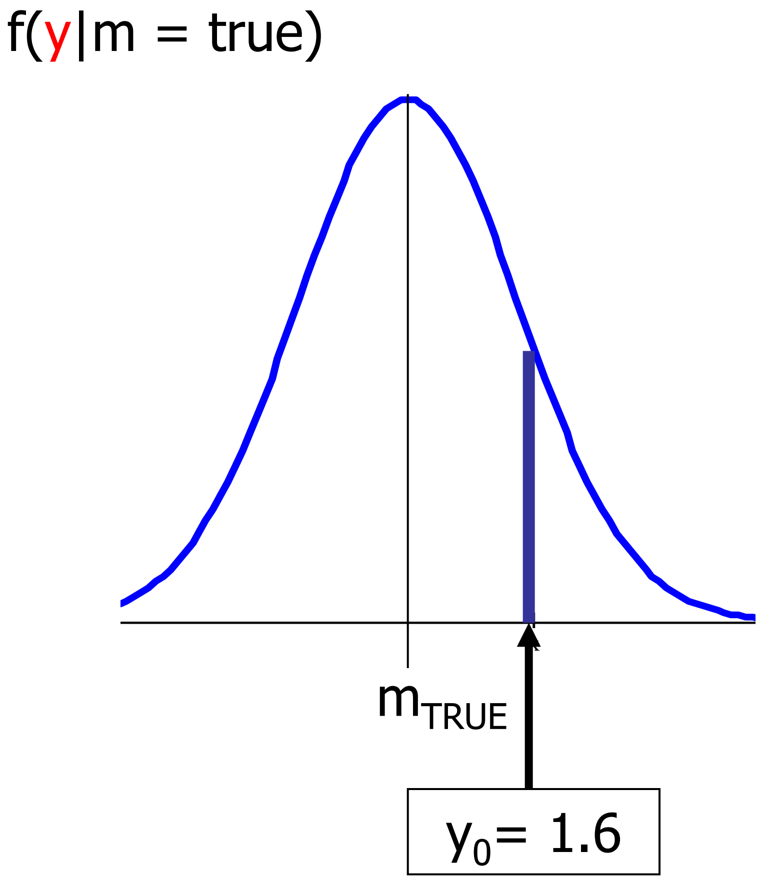
Likelihood



Likelihood



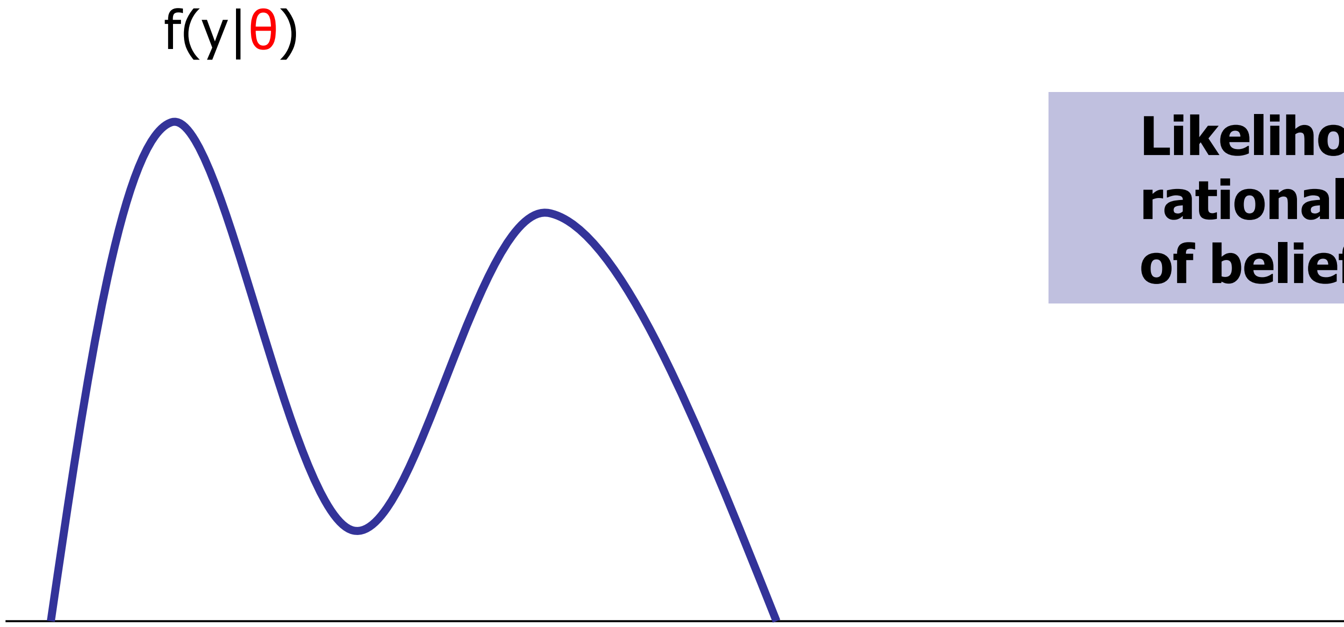
Maximum Likelihood



The true value does not maximise the probability of the sample!

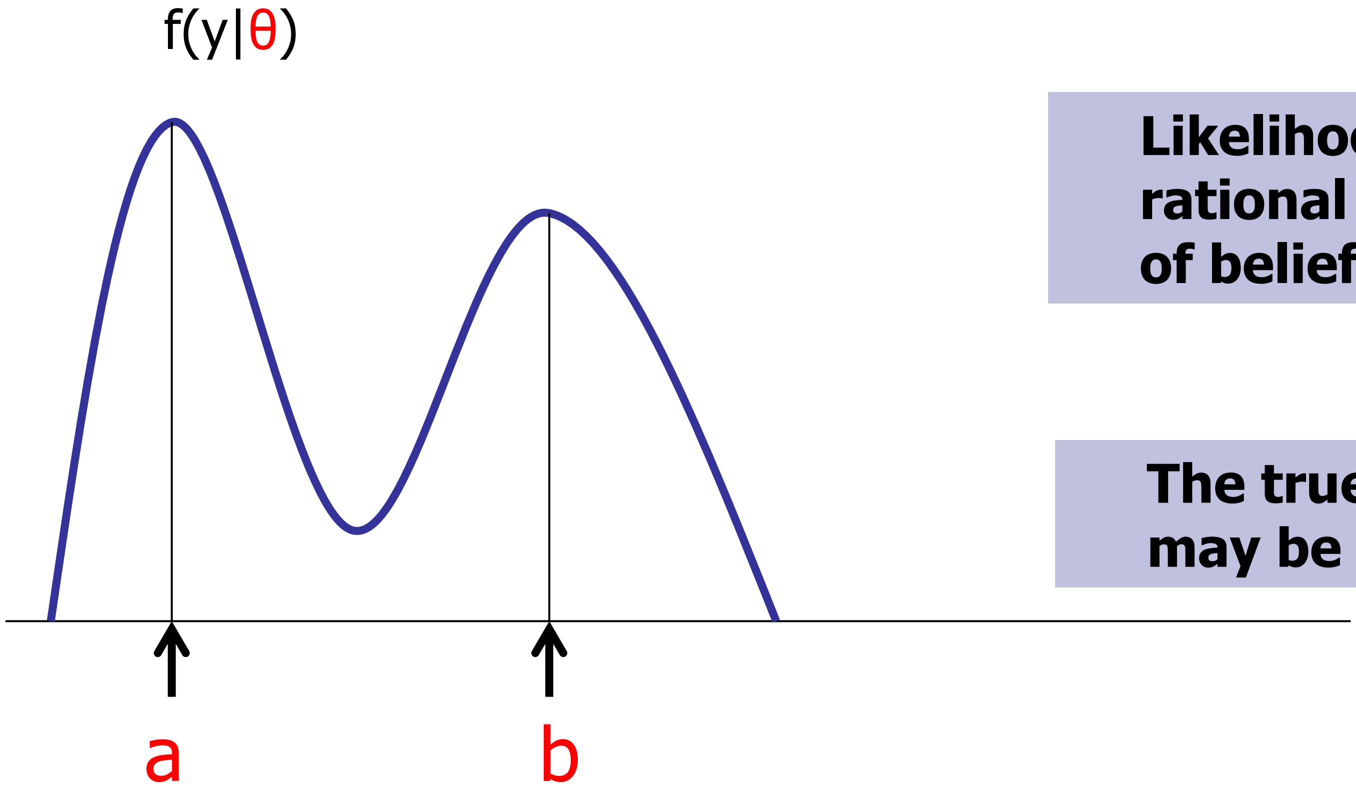
ML would maximise the probability of the sample only if the parameter were the true value.

Maximum Likelihood



Likelihood is a rational degree of belief.

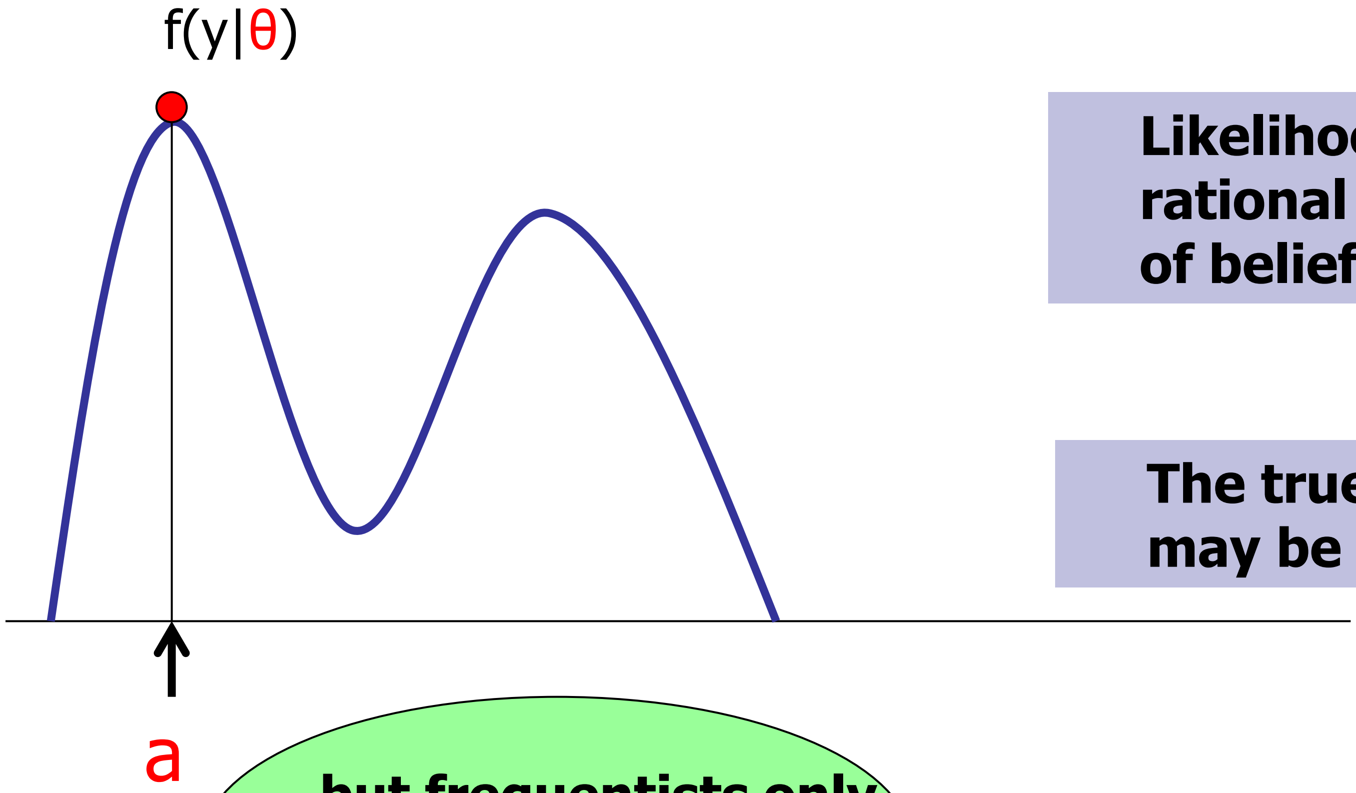
Maximum Likelihood



Likelihood is a rational degree of belief.

The true value may be a or b.

Maximum Likelihood

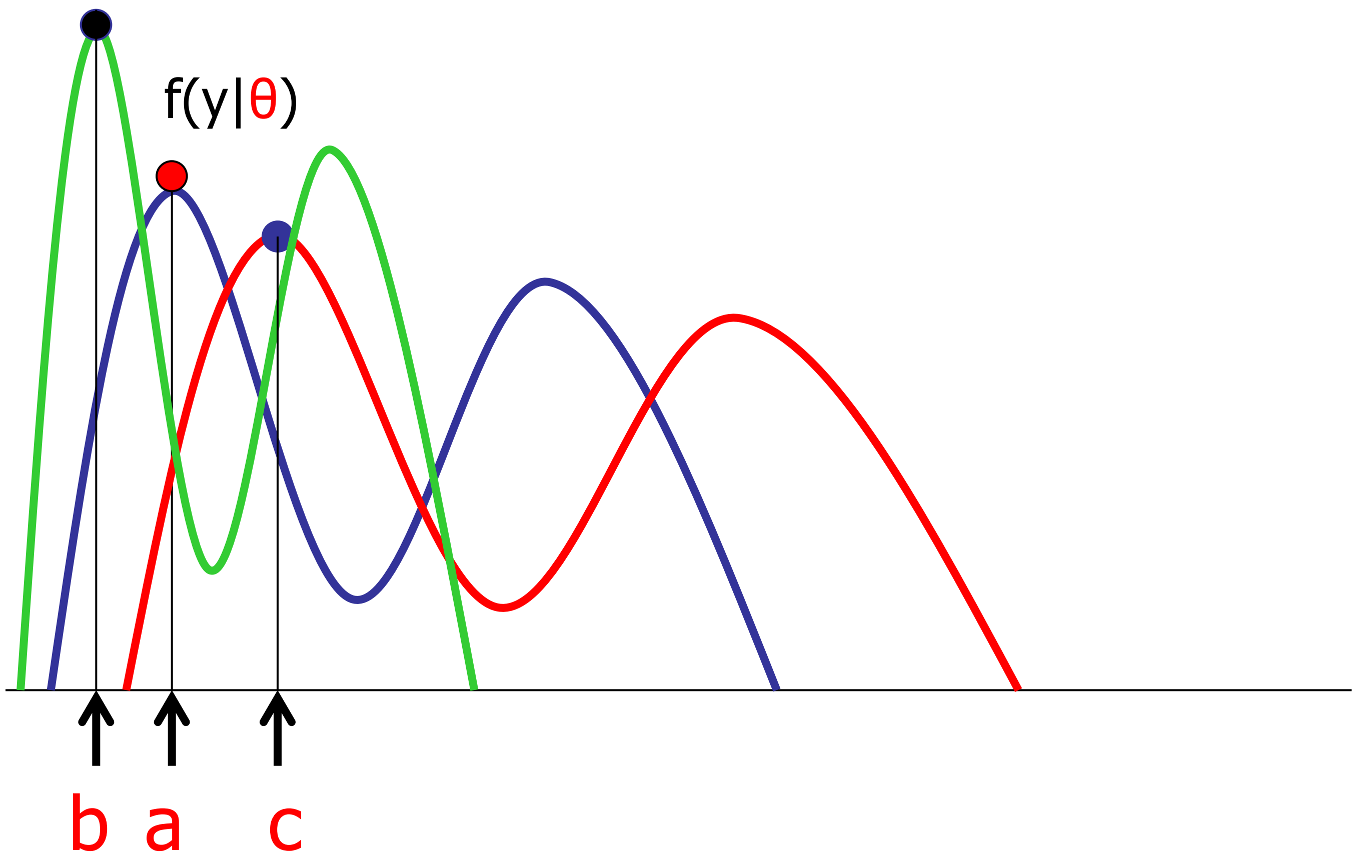


Likelihood is a rational degree of belief.

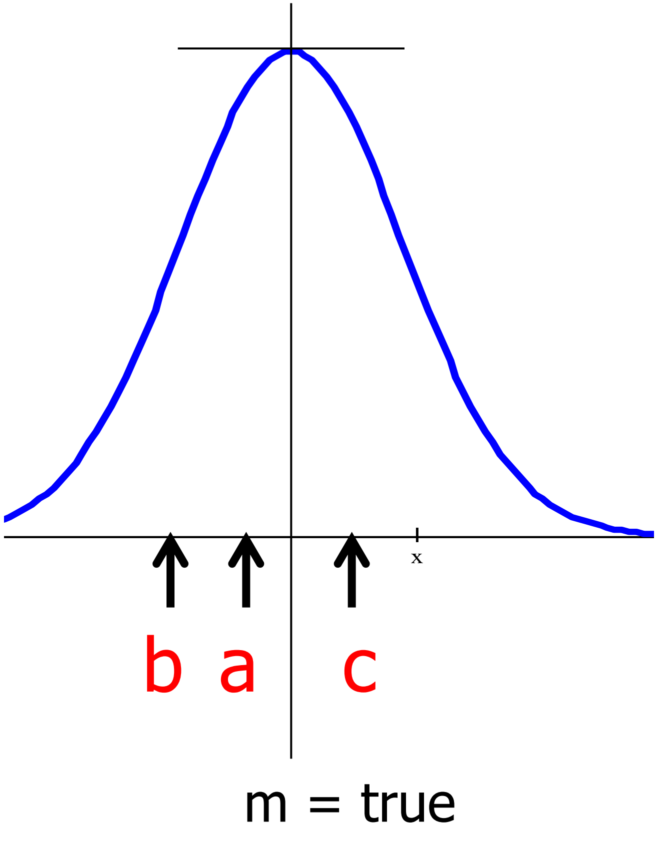
The true value may be a OR b.

… but frequentists only consider the maximum!

Maximum Likelihood



Maximum Likelihood



Maximum has favourable distributional properties

...but only if n is sufficiently large **!!**

Maximum Likelihood

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{e}$$

$$f\left(\mathbf{y} \mid \mu, \sigma^2\right) = \frac{1}{\left(2\pi\right)^{\frac{n}{2}}\left(\sigma^2\right)^{\frac{n}{2}}} \exp\left[\sum_1^n -\frac{\left(y_i - \mu\right)^2}{2\sigma^2}\right]$$

$$\log f = cte - \frac{n}{2}\log\left(\sigma^2\right) - \frac{1}{2\sigma^2}\left[\sum_1^n \left(y_i - \mu\right)^2\right]$$

Maximum Likelihood

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{e}$$

$$\log f = cte - \frac{n}{2}\log\left(\sigma^2\right) - \frac{1}{2\sigma^2}\left[\sum_1^n \left(y_i - \mu\right)^2\right]$$

$$\frac{\partial}{\partial \sigma^2}\log f = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4}\left[\sum_1^n \left(y_i - \mu\right)^2\right] = 0 \quad \rightarrow \quad \hat{\sigma}^2 = \frac{1}{n}\sum_1^n \left(y_i - \mu\right)^2$$

$$\frac{\partial}{\partial \mu}\log f = \frac{1}{2\sigma^2}\left[2\sum_1^n \left(y_i - \mu\right)\right] = \sum_1^n \left(y_i\right) + n\mu = 0 \quad \rightarrow \quad \hat{\mu} = \frac{1}{n}\sum_1^n \left(y_i\right)$$

Maximum Likelihood

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{e}$$

$$\hat{\mu} = \frac{1}{n}\sum_{1}^{n}\left(y_i\right)$$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{1}^{n}\left(y_i - \hat{\mu}\right)^2$$

REML

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{e}$$

$$y_i^* = y_i - y_k = \mu + e_i - \left(\mu + e_k\right) = e_i - e_k$$

REML

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{e}$$

$$y_i^* = y_i - y_k = \mu + e_i - \left(\mu + e_k\right) = e_i - e_k$$

$$\mathbf{K}'\mathbf{y} = \begin{bmatrix} 1 & -1 & 0 & 0 & \dots & \dots & 0 \\ 1 & 0 & -1 & 0 & \dots & \dots & 0 \\ 1 & 0 & 0 & -1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & 0 & 0 & 0 & \dots & \dots & -1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix}$$

REML

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{e}$$

$$y_i^* = y_i - y_k = \mu + e_i - \left(\mu + e_k\right) = e_i - e_k$$

$$\mathbf{K}'\mathbf{y} = \begin{bmatrix} 1 & -1 & 0 & 0 & \dots & \dots & 0 \\ 1 & 0 & -1 & 0 & \dots & \dots & 0 \\ 1 & 0 & 0 & -1 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 1 & 0 & 0 & 0 & \dots & \dots & -1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} y_1 - y_2 \\ y_1 - y_3 \\ y_1 - y_4 \\ \vdots \\ y_1 - y_n \end{bmatrix} = \begin{bmatrix} y_1^* \\ y_2^* \\ y_3^* \\ \vdots \\ y_{n-1}^* \end{bmatrix} = \mathbf{y}^*$$

There are *n* data *y* but only *n-1* data *y**
The space now possesses *n-1* dimensions
One degree of freedom has been lost

REML

$$\mathbf{y} = \mathbf{1}\mu + \mathbf{e}$$

$$\hat{\sigma}^2_{REML} = \frac{1}{n-1}\left(\mathbf{y} - \mathbf{1}\frac{1}{n}\sum_1^n y_i\right)'\left(\mathbf{y} - \mathbf{1}\frac{1}{n}\sum_1^n y_i\right)$$

$$\hat{\sigma}^2_{REML} = \frac{1}{n-1}(\mathbf{y} - \mathbf{1}\hat{\mu})'(\mathbf{y} - \mathbf{1}\hat{\mu})$$

REML

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e}$$

$$\hat{\sigma}^2_{REML} = \frac{1}{n-1}\left(\mathbf{y} - \mathbf{1}\hat{\mu}\right)'\left(\mathbf{y} - \mathbf{1}\hat{\mu}\right)$$

$$\hat{\sigma}^2_{REML} = \frac{1}{n - \mathrm{rg}\left(\mathbf{X}\right)}\left(\mathbf{y} - \mathbf{X}\hat{\mathbf{b}}\right)'\left(\mathbf{y} - \mathbf{X}\hat{\mathbf{b}}\right)$$

Maximum Likelihood

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{Z}\mathbf{a} + \mathbf{e}$$

$$\left[\mathrm{var}\left(\sigma_A^2\right) \quad \mathrm{var}\left(\sigma_e^2\right)\right] \quad \rightarrow \quad \mathrm{var}\left(h^2\right) = \mathrm{var}\left(\frac{\sigma_A^2}{\sigma_A^2 + \sigma_e^2}\right)$$

Maximum Likelihood

$$f\left(x,y\right) \approx \left[f\left(x,y\right)\right]_{\substack{x_0\\y_0}} + \left[\frac{\partial f}{\partial x}\right]_{\substack{x_0\\y_0}}\left(x - x_0\right) + \left[\frac{\partial f}{\partial y}\right]_{\substack{x_0\\y_0}}\left(y - y_0\right)$$

$$h^2 = \frac{\sigma_A^2}{\sigma_A^2 + \sigma_e^2} \approx \left[h^2\right]_{\substack{\sigma_A^2 = \hat{\sigma}_A^2\\ \sigma_e^2 = \hat{\sigma}_e^2}} + \left[\frac{\partial h^2}{\partial \sigma_A^2}\right]_{\substack{\sigma_A^2 = \hat{\sigma}_A^2\\ \sigma_e^2 = \hat{\sigma}_e^2}}\left(\sigma_A^2 - \hat{\sigma}_A^2\right) + \left[\frac{\partial h^2}{\partial \sigma_e^2}\right]_{\substack{\sigma_A^2 = \hat{\sigma}_A^2\\ \sigma_e^2 = \hat{\sigma}_e^2}}\left(\sigma_e^2 - \hat{\sigma}_e^2\right)$$

$$= \frac{\hat{\sigma}_A^2}{\hat{\sigma}_A^2 + \hat{\sigma}_e^2} + k_1\left(\sigma_A^2 - \hat{\sigma}_A^2\right) + k_2\left(\sigma_e^2 - \hat{\sigma}_e^2\right)$$

Maximum Likelihood

ML is asymptotically unbiased

$$E\left(\hat{\sigma}_A^2\right) = \sigma_A^2 \qquad\qquad E\left(\hat{\sigma}_e^2\right) = \sigma_e^2$$

$$h^2 \approx \quad = \frac{\hat{\sigma}_A^2}{\hat{\sigma}_A^2 + \hat{\sigma}_e^2} + k_1\left(\sigma_A^2 - \hat{\sigma}_A^2\right) + k_2\left(\sigma_e^2 - \hat{\sigma}_e^2\right)$$

$$\hat{h}^2 = E\left(h^2\right) \approx \left[h^2\right]_{\substack{\sigma_A^2 = \hat{\sigma}_A^2 \\ \sigma_e^2 = \hat{\sigma}_e^2}} + 0 + 0 = \frac{\hat{\sigma}_A^2}{\hat{\sigma}_A^2 + \hat{\sigma}_e^2}$$

Maximum Likelihood

$$\text{var}\left(\hat{h}^2\right) = E\left[h^2 - E\left(h^2\right)\right]^2 \approx E\left[k_1\left(\sigma_A^2 - \hat{\sigma}_A^2\right) + k_2\left(\sigma_e^2 - \hat{\sigma}_e^2\right)\right]^2 =$$

$$= k_1^2 \,\text{var}\left(\sigma_A^2\right) + k_2^2 \,\text{var}\left(\sigma_e^2\right)$$

$$s.e.\left(\hat{h}^2\right) = \sqrt{\text{var}\left(\hat{h}^2\right)}$$