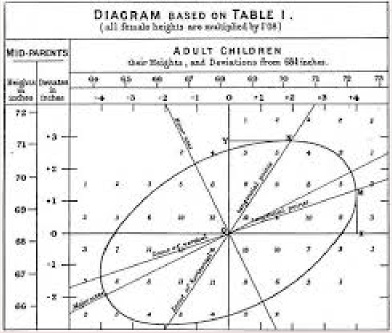
Chapter 4

Linear model



- Origin: Guass, Galton.

Why Normal?

- The normal distribution, is the most important distribution in statistics.
- Many mechanism creates an end product that follows a Normal distribution. Like sums of random variables, products of random variables.
- The distribution is the easiest to handle computationally.

Sum of Uniform

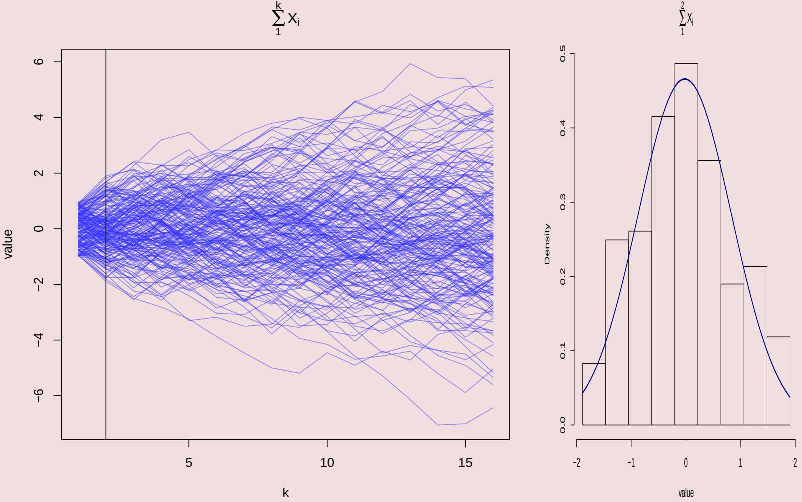
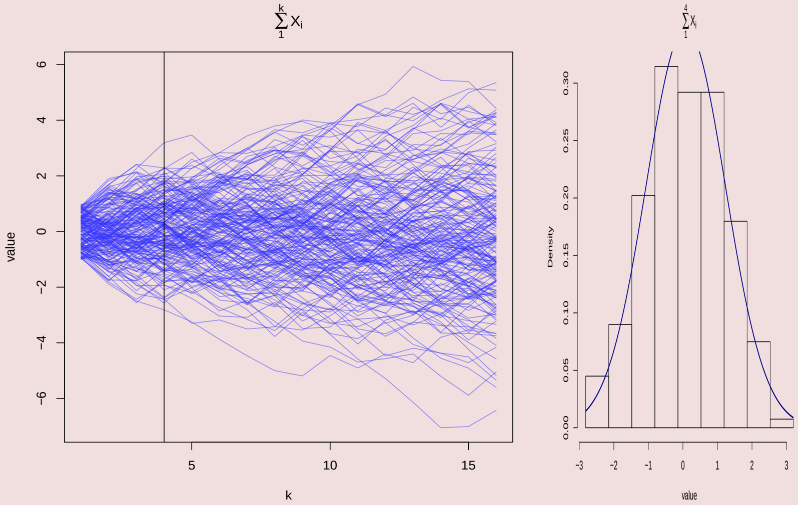


Figure : $\sum^k$ X $p(x) = 1$ (x)

Sum of Uniform



Sum of Uniform

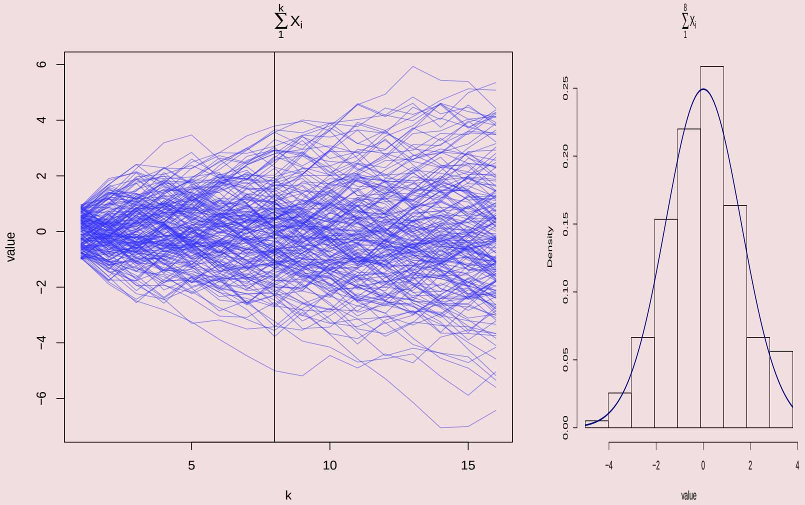


Figure : $\sum^{k} \quad X \quad p(x) = \frac{1}{} \mathbb{I} \quad (x)$

Sum of Uniform

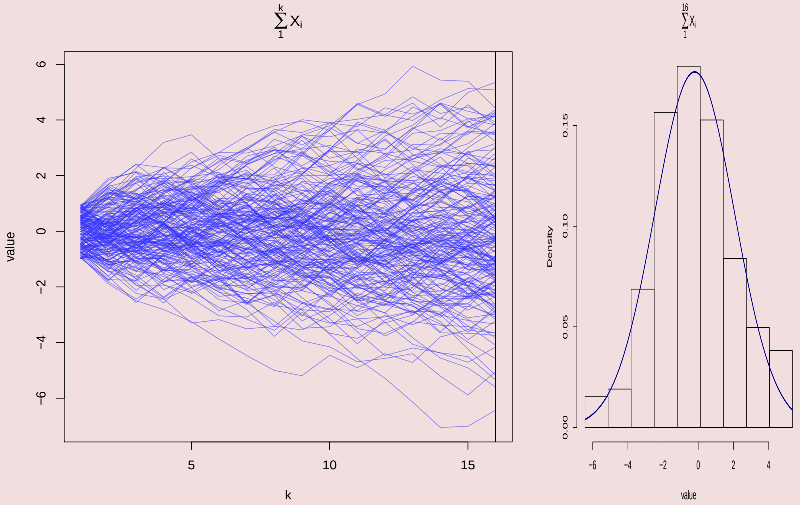


Figure : $\sum^k \quad X \quad p(x) = \frac{1}{} \mathbb{I} \quad (x)$

Sum of Poisson

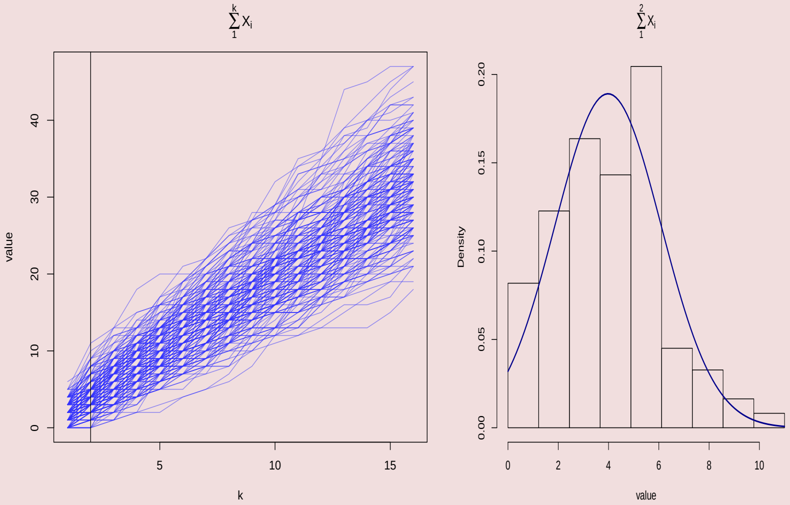


Figure : $\sum^{k} \; X \; p(x|\lambda) = \lambda^{x} e^{-\lambda}$

Sum of Poisson

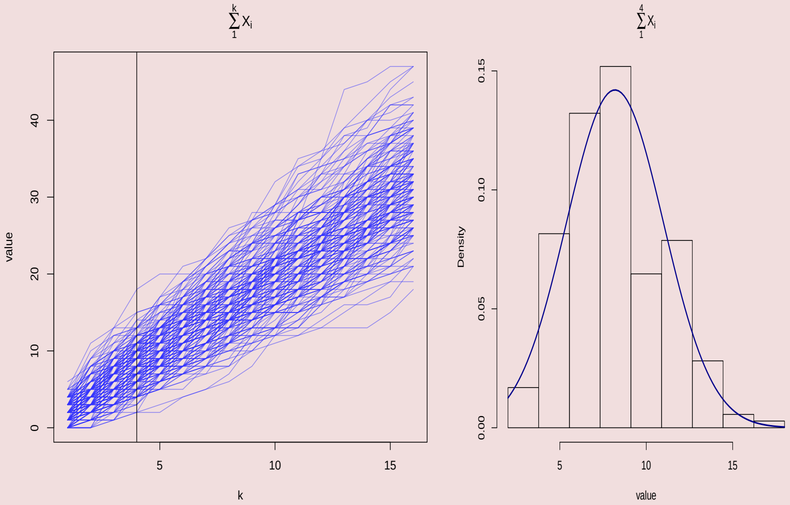


Figure : $\sum^k X$ ... $p(x|\lambda) = \lambda^x e^{-\lambda}$

Sum of Poisson

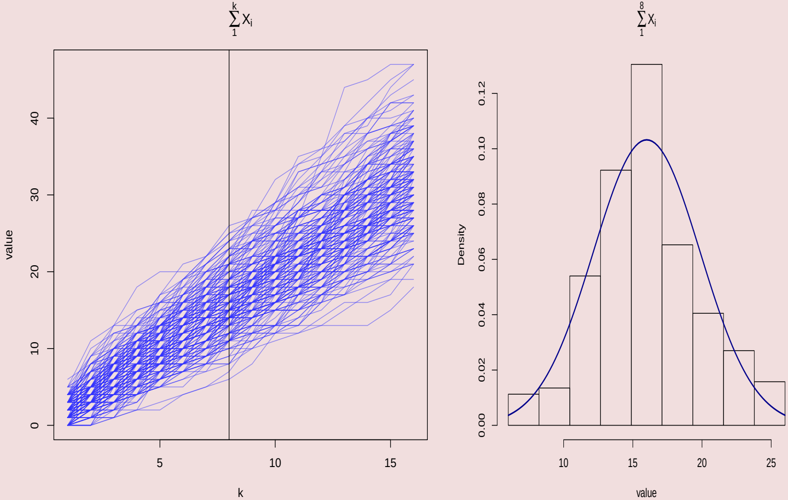


Figure : $\sum^k \quad X \quad p(x|\lambda) = \lambda^x e^{-\lambda}$

Sum of Poisson

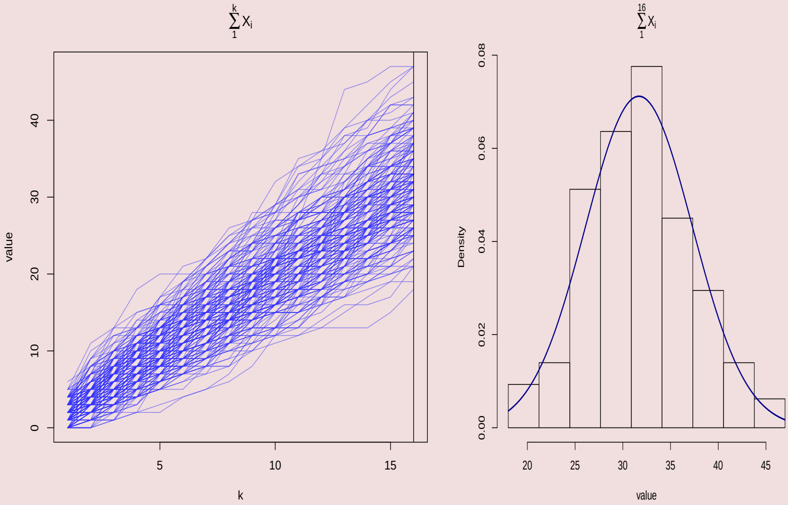


Figure : $\sum^k \; X \quad p(x|\lambda) = \lambda^x e^{-\lambda}$

Model

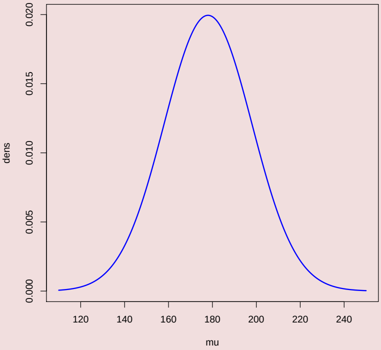
$$n_F \sim Binomial(n, p)$$
$$p \sim U[0, 1]$$

A first model of height

$$
\begin{aligned}
h_i &\sim N(\mu, \sigma) \\
\mu &\sim N(178, 20) \\
\sigma &\sim U[0, 50]
\end{aligned}
$$

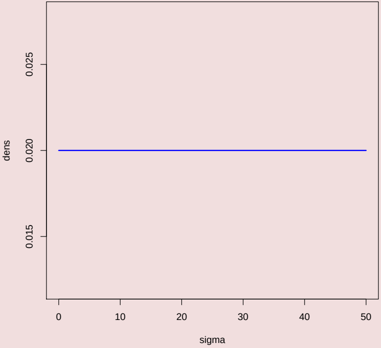
A first model of height

$$h_i \sim N(\mu, \sigma)$$
$$\mu \sim N(178, 20)$$
$$\sigma \sim U[0, 50]$$



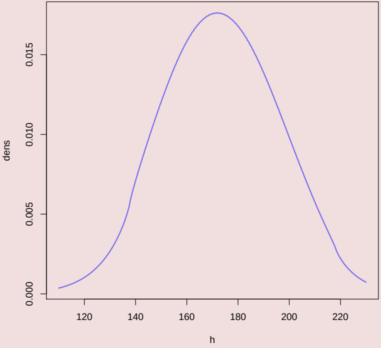
A first model of height

$$h_i \sim N(\mu, \sigma)$$
$$\mu \sim N(178, 20)$$
$$\sigma \sim U[0, 50]$$



A first model of height

$$h_i \sim N(\mu, \sigma)$$
$$\mu \sim N(178, 20)$$
$$\sigma \sim U[0, 50]$$



Model to distribution

Model	density
$h_i \sim N(\mu, \sigma)$	$p(h_i \vert \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2\sigma^2}(h_i - \mu_i)^2}$
$\mu \sim N(178, 20)$	$p(\beta) = \frac{1}{\sqrt{2\pi 20}} e^{-\frac{1}{20^2}(\beta - 178)^2}$
$\sigma \sim U(0, 50)$	$p(\sigma) = \frac{1}{50} \mathbb{I}_{[0,50]}(\sigma)$

Unless stated variables are independent.

Posterior distribution of μ

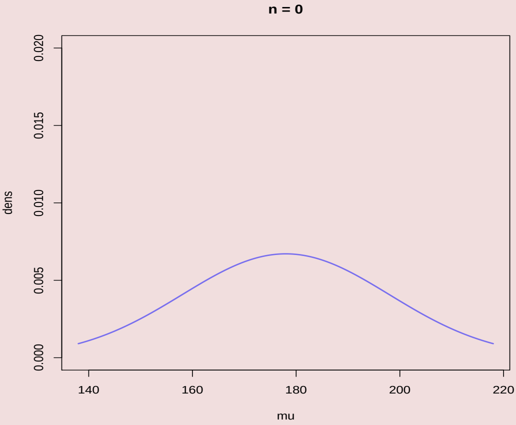


Figure : $p(\mu | h_1, \ldots, h_n)$

Posterior distribution of μ

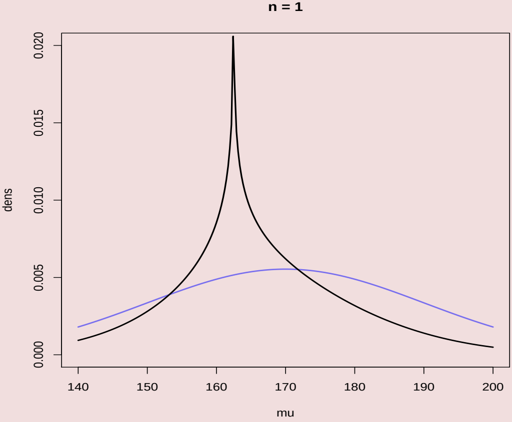


Figure : $p(\mu|h_1, \dots, h_n)$

Posterior distribution of μ

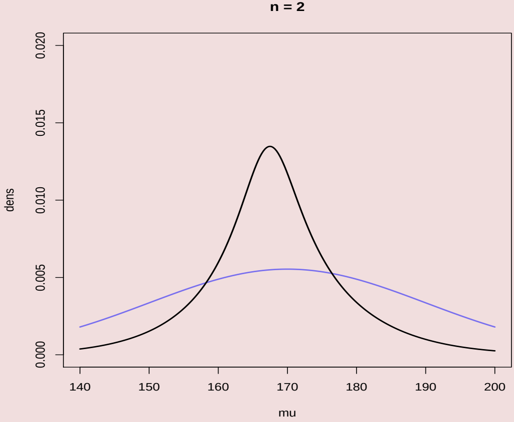


Figure : $p(\mu|h_1, \ldots, h_n)$

Posterior distribution of μ

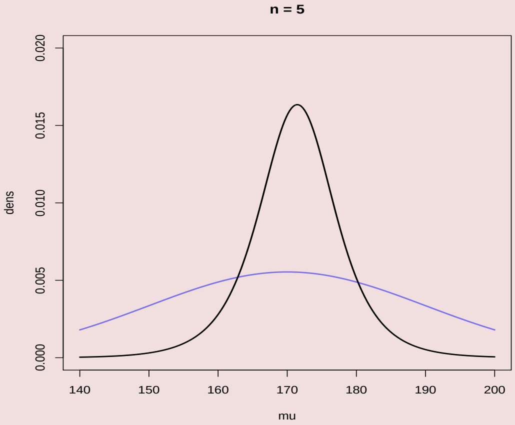


Figure : $p(\mu|h_1,\ldots,h_n)$

Posterior distribution of μ

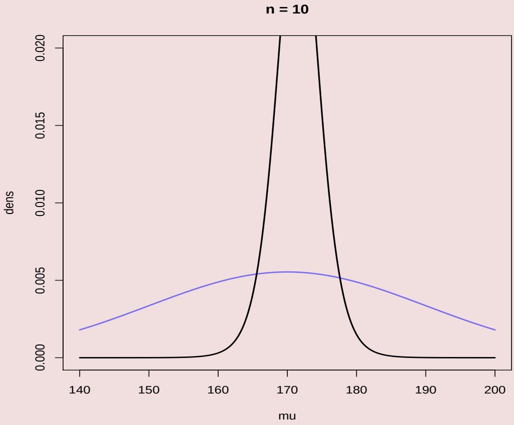


Figure : $p(\mu|h_1, \ldots, h_n)$

Predictive distribution, $p(y|\cdot)$

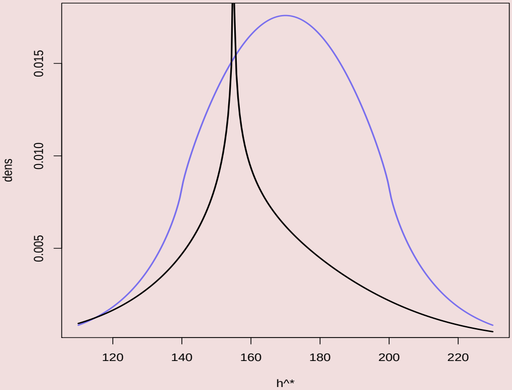


Figure : $p(h^*|h_1, \ldots, h_n)$

Predictive distribution, $p(y|\cdot)$

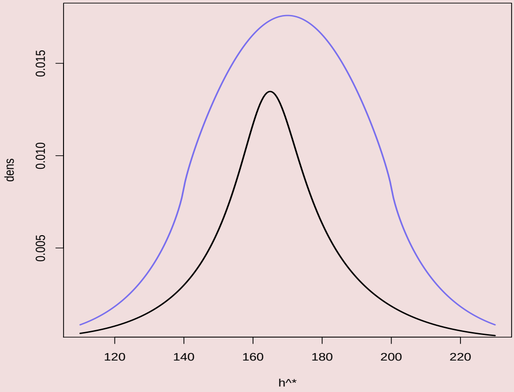


Figure : $p(h^*|h_1, \ldots, h_n)$

Predictive distribution, $p(y|\cdot)$

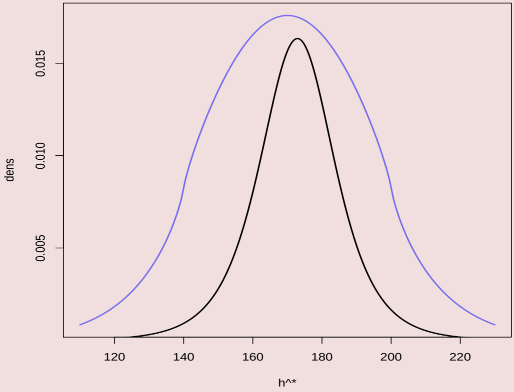


Figure : $p(h^*|h_1,\ldots,h_n)$

Predictive distribution, $p(y|\cdot)$

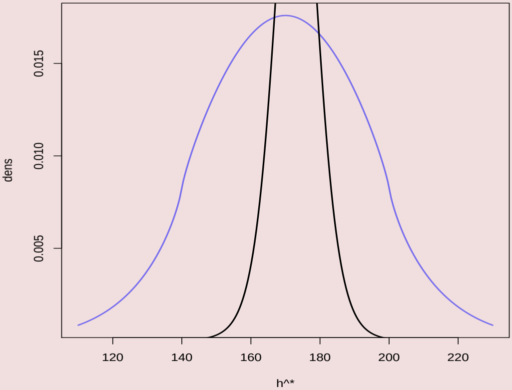


Figure : $p(h^*|h_1, \ldots, h_n)$

Predictive distribution, $p(y|\cdot)$

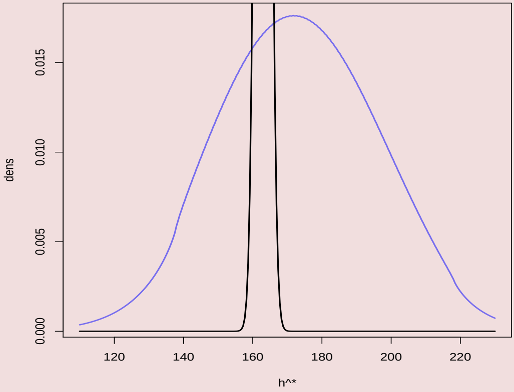


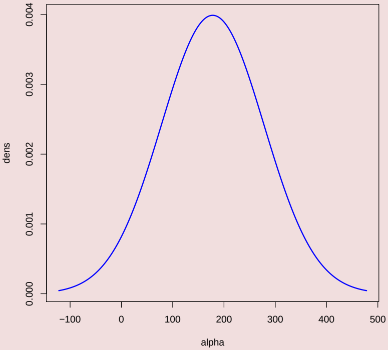
Figure : $p(h^*|h_1,\ldots,h_n)$

Adding a predictor

$$
\begin{aligned}
h_i &\sim N(\mu_i, \sigma) \\
\mu_i &= \alpha + w_i\beta \\
\alpha &\sim N(178, 100) \\
\beta &\sim N(0, 20) \\
\sigma &\sim U[0, 50]
\end{aligned}
$$

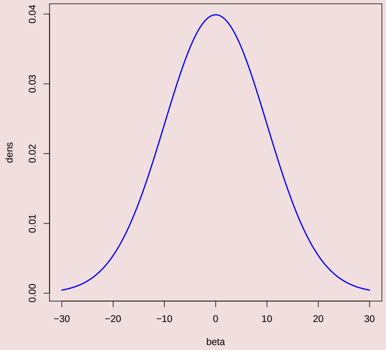
Adding a predictor

$$
\begin{aligned}
h_i &\sim N(\mu_i, \sigma) \\
\mu_i &= \alpha + w_i \beta \\
\alpha &\sim N(178, 100) \\
\beta &\sim N(0, 20) \\
\sigma &\sim U[0, 50]
\end{aligned}
$$



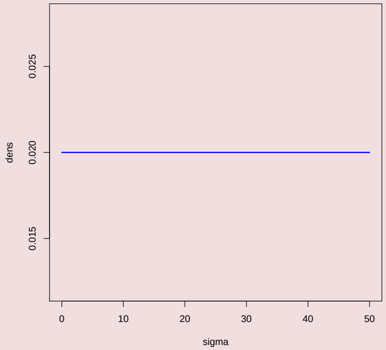
Adding a predictor

$$
\begin{aligned}
h_i &\sim N(\mu_i, \sigma) \\
\mu_i &= \alpha + w_i \beta \\
\alpha &\sim N(178, 100) \\
\beta &\sim N(0, 20) \\
\sigma &\sim U[0, 50]
\end{aligned}
$$



Adding a predictor

$$h_i \sim N(\mu_i, \sigma)$$
$$\mu_i = \alpha + w_i\beta$$
$$\alpha \sim N(178, 100)$$
$$\beta \sim N(0, 20)$$
$$\sigma \sim U[0, 50]$$



Model to distribution

Model	density
$h_i \sim N(\mu_i, \sigma)$	$p(h_i \mid \mu_i, \sigma) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2\sigma^2}(h_i - \mu_i)^2}$
$\mu_i = w_i \cdot \beta$	
$\beta \sim N(0, 20)$	$p(\beta) = \frac{1}{\sqrt{2\pi 20}} e^{-\frac{1}{20^2}(\beta)^2}$
$\sigma \sim U(0, 1)$	$p(\sigma) = \frac{1}{50} \mathbb{I}_{[0,50]}(\sigma)$

Unless stated variables are independent.

Predictive function $\mu(x)$

- In the book, weight is from 31 to 63.
- We have a prior on the function:

$$\mu(x) = \alpha + x\beta$$

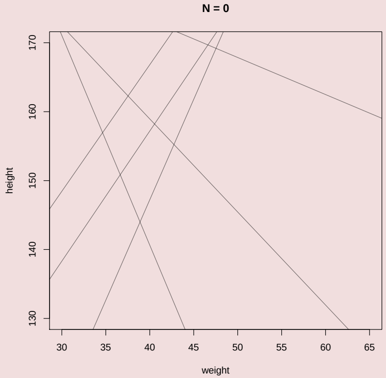


Figure : Prior draws of $\mu(x)$.

Posterior $\mu(x)$

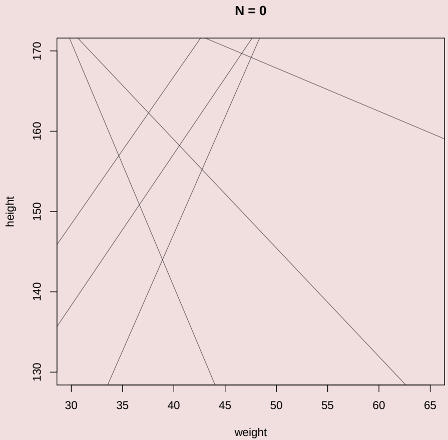


Figure : $\mu(x) = \alpha + x\beta | h_1, \ldots, h_n$

Posterior $\mu(x)$

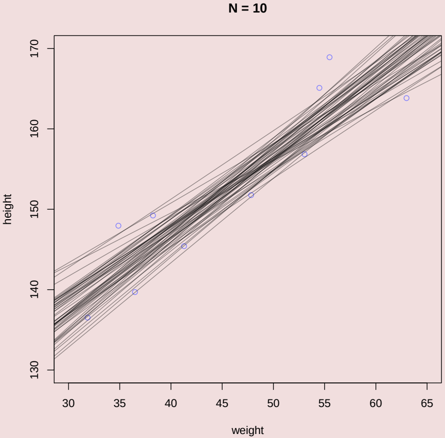


Figure : $\mu(x) = \alpha + x\beta | h_1, \ldots, h_n$

Posterior $\mu(x)$

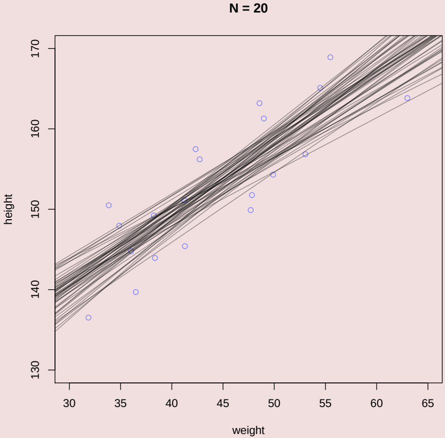


Figure : $\mu(x) = \alpha + x\beta | h_1, \ldots, h_n$

Posterior $\mu(x)$

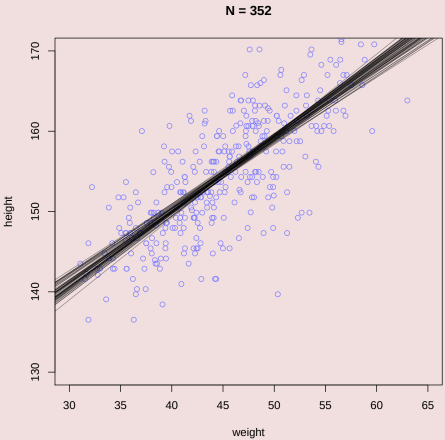


Figure : $\mu(x) = \alpha + x\beta | h_1, \ldots, h_n$

Posterior $\mu(x)$

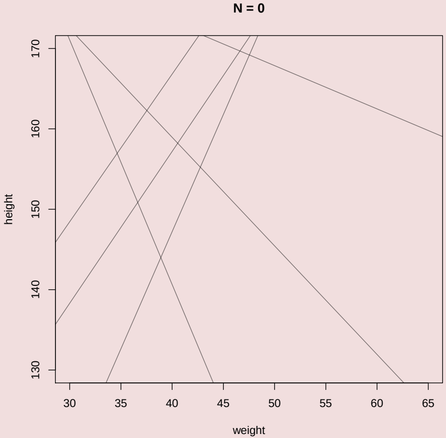


Figure : $\mu(x) = \alpha + x\beta | h_1, \ldots, h_n$

Posterior $\mu(x)$

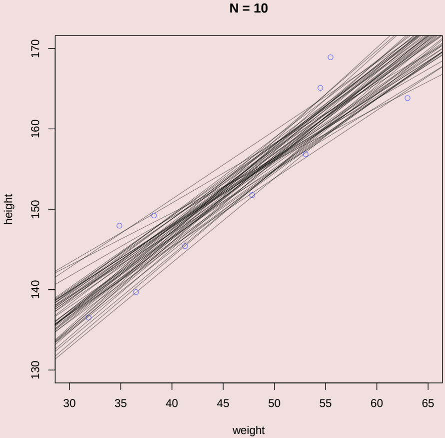


Figure : $\mu(x) = \alpha + x\beta | h_1, \ldots, h_n$

Posterior $\mu(x)$

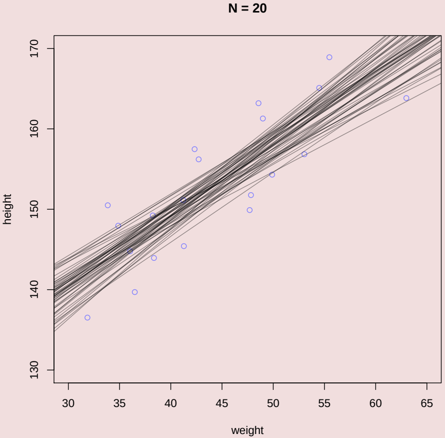


Figure : $\mu(x) = \alpha + x\beta | h_1, \ldots, h_n$

Posterior $\mu(x)$

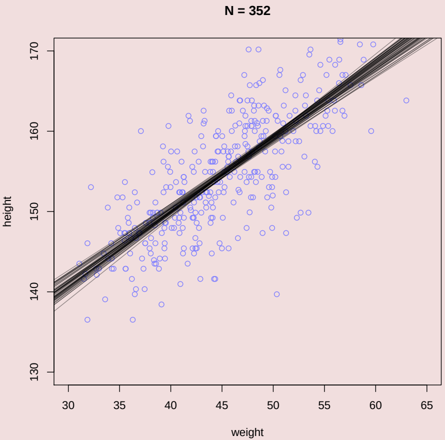


Figure : $\mu(x) = \alpha + x\beta | h_1, \ldots, h_n$

Simple Height model

$$h_i \sim N(\mu, \sigma)$$
$$\mu \sim N(178, 20)$$
$$\sigma \sim U[0, 50]$$

```
library(rethinking)
data(Howell1)
dataHeight <- Howell1[Howell1$age >= 18,]

# building the model
model <- map(
             flist = alist(
               height ~ dnorm(mu, sigma),
               mu     ~ dnorm(178,20),
               sigma  ~ dunif(0,50)
             ),
             data  = dataHeight)
```

predictive Height model

$$h_i \sim N(\mu_i, \sigma)$$
$$\mu_i = \alpha + x_i\beta$$
$$\alpha \sim N(178, 100)$$
$$\beta \sim N(0, 20)$$
$$\sigma \sim U[0, 50]$$

```
library(rethinking)
data(Howell1)
dataHeight <- Howell1[Howell1$age >= 18,]

# building the model
model2 <- map(
flist = alist(
        height ~ dnorm(mu, sigma),
        mu     <- alpha + weight * beta ,
        alpha  ~ dnorm(156, 100),
        beta   ~ dnorm(0  , 10),
        sigma  ~ dunif(0,50)
),
data   = dataHeight)
```

Posterior Samples from the models

```
post <- extract.samples(model, n = 100)
head(post, n = 3)
```

Output from the first model:

	mu	sigma
1	154.6512	7.950005
2	154.2872	7.950664
3	154.1929	7.704647

Output from the second model:

	alpha	beta	sigma
1	115.8987	0.8559246	5.170375
2	114.2855	0.8911623	5.181440
3	111.4015	0.9479790	5.143857

Comparing the models

The posterior distribution for h^* given $x^* = 30, 50$:

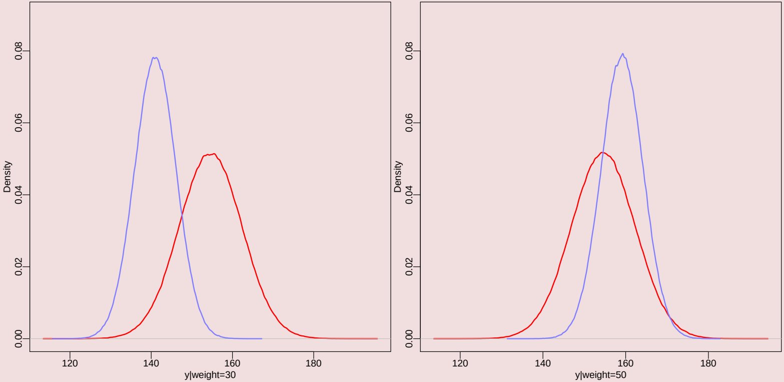


Figure : $p(y^*|x^* = \cdot, h_1, \ldots, h_n)$, blue predictive model and red simple model.