



Bayesian Statistics and Hierarchical Bayesian Modeling for Psychological Science

Lecture 06

Lei Zhang

Social, Cognitive and Affective Neuroscience Unit (SCAN-Unit)
Department of Basic Psychological Research and Research Methods

https://github.com/lei-zhang/BayesCog_Wien

lei.zhang@univie.ac.at
lei-zhang.net
 @lei_zhang_lz



universität
wien
Fakultät für Psychologie

Evaluate the review

- the ability to briefly and clearly summarize the main findings
- critical thinking
 - regarding the design
 - regarding the analyses (stats)
- appreciate the use of cognitive modeling
 - is the modeling approach appropriate to answer the research question?
 - is the interpretation sound?
 - could there be alternative models?
 - model recovery
 - posterior check etc.

1 st	2 nd
70%	40%
30%	60%

Recap

What we've learned...

- R Basics
- probability distributions
- Bayes' theorem, $p(\theta|D)$
- Binomial model
- MCMC and Stan

LINEAR REGRESSION

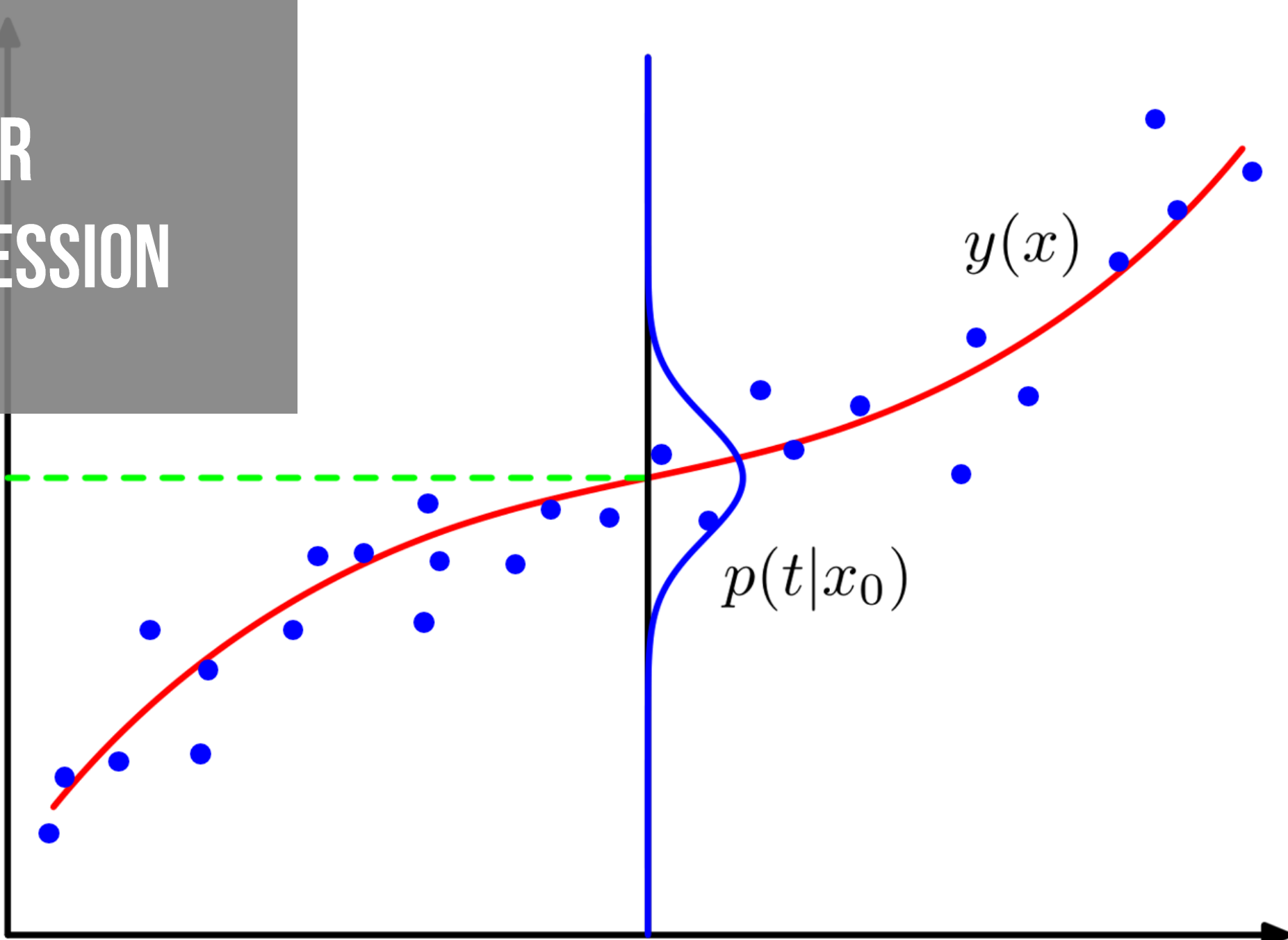
$y(x_0)$

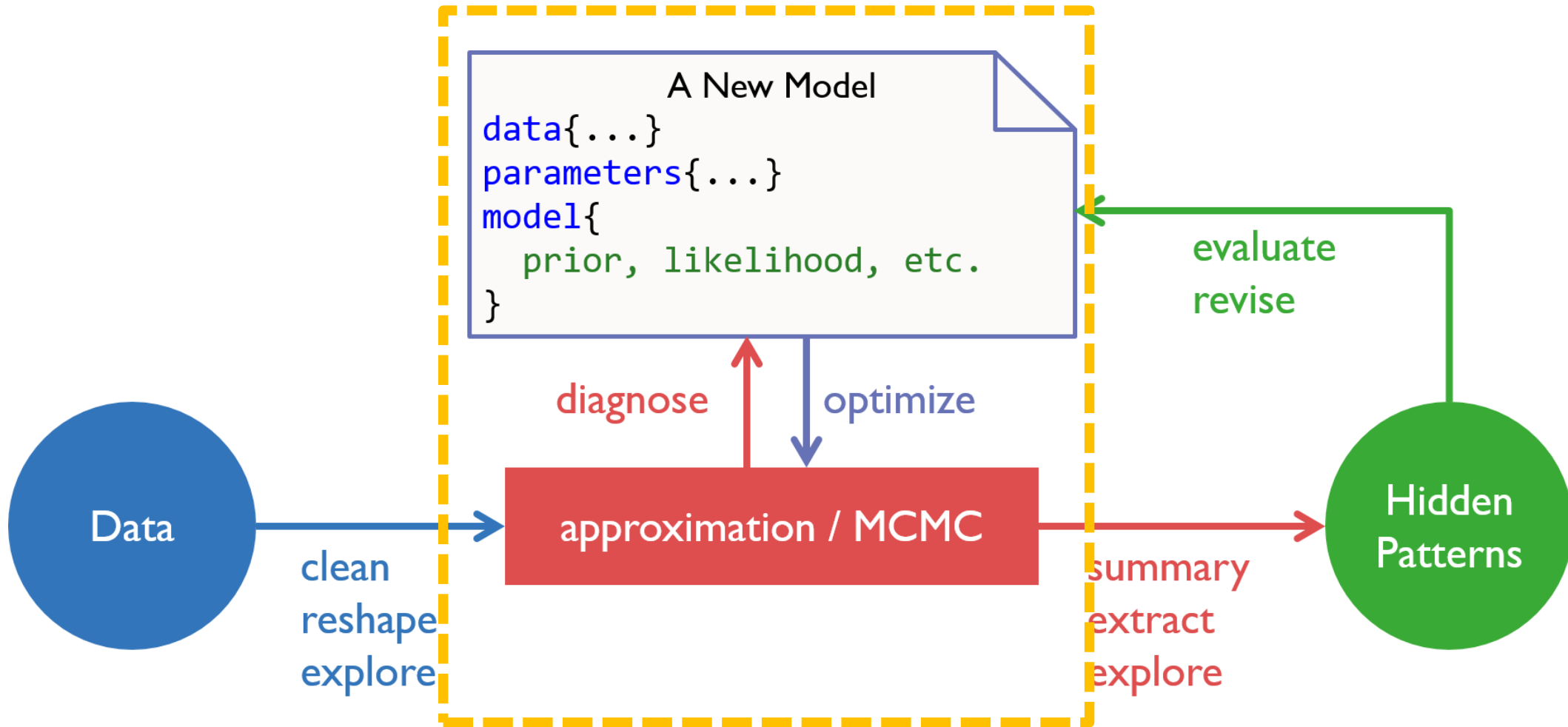
$p(t|x_0)$

$y(x)$

x_0

x





Linear Regression: height ~ weight

cognitive model

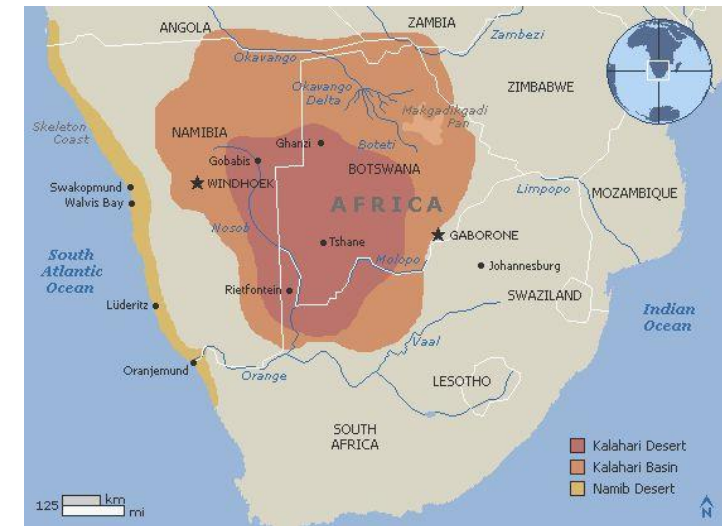
statistics

computing

```
.../04.regression_height/_scripts/regression_height_main.R
```

make scatter plot and fit the model with `lm()`

```
>load('_data/height.RData')
>d <- Howell1
>d <- d[ d$age >= 18 , ]
>head(d)
height  weight age male
1 151.765 47.82561 63    1
2 139.700 36.48581 63    0
3 136.525 31.86484 65    0
4 156.845 53.04191 41    1
5 145.415 41.27687 51    0
6 163.830 62.99259 35    1
```



Results with lm()

cognitive model

statistics

computing

```
> L <- lm( height ~ weight, d) # estimate model by minimizing least squares errors
> summary(L)
```

Call:

```
lm(formula = height ~ weight, data = d)
```

Residuals:

Min	1Q	Median	3Q	Max
-19.7464	-2.8835	0.0222	3.1424	14.7744

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)	
(Intercept)	113.87939	1.91107	59.59	<2e-16	***
weight	0.90503	0.04205	21.52	<2e-16	***

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 5.086 on 350 degrees of freedom

Multiple R-squared: 0.5696, Adjusted R-squared: 0.5684

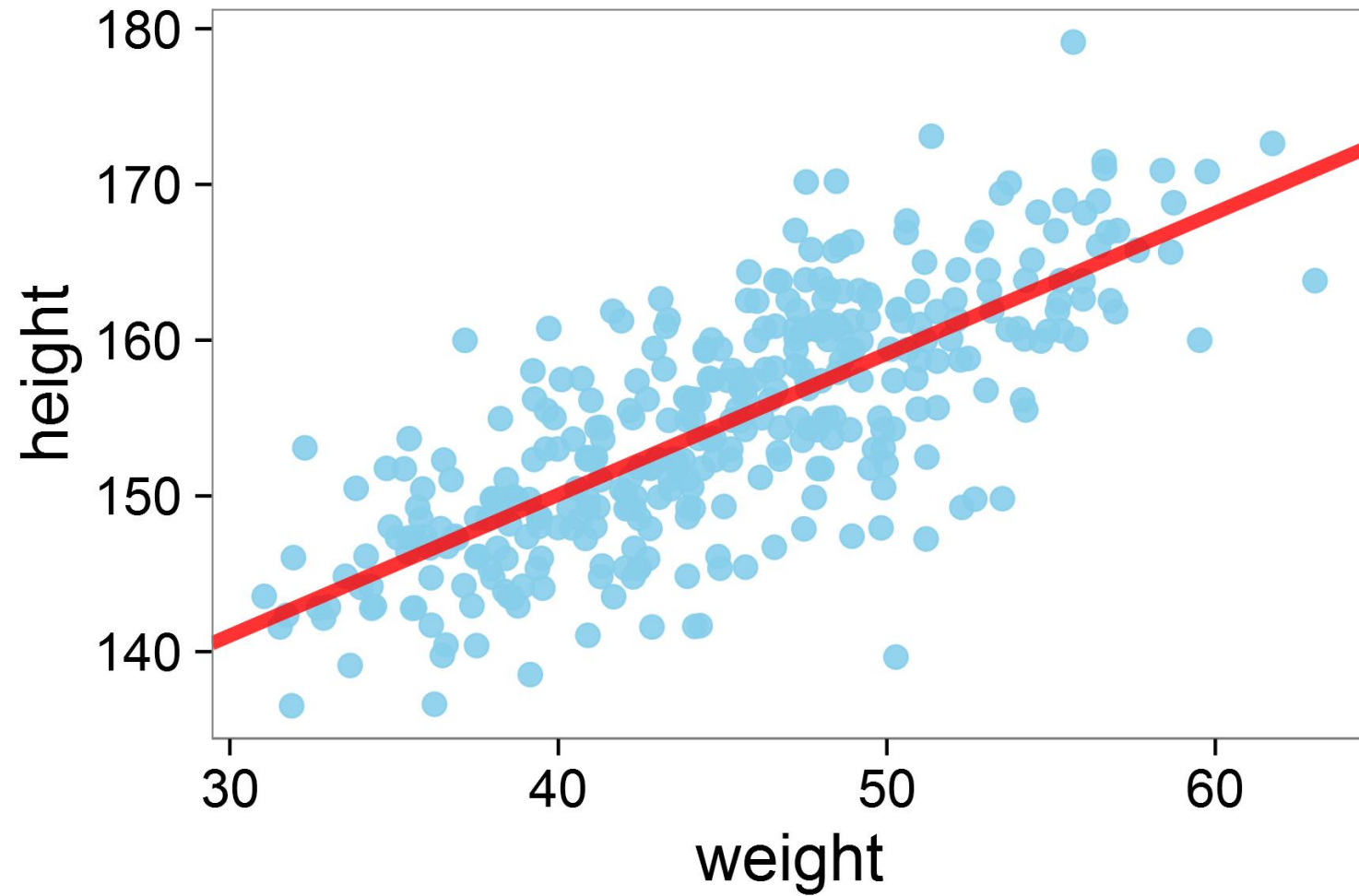
F-statistic: 463.3 on 1 and 350 DF, p-value: < 2.2e-16

height ~ weight

cognitive model

statistics

computing



Rethinking Regression Model

cognitive model

statistics

computing

$$\mu_i = \alpha + \beta x_i$$

~~$$y_i = \mu_i + \varepsilon$$~~

~~$$\varepsilon \sim \text{Normal}(0, \sigma)$$~~

$$y_i \sim \text{Normal}(\mu_i, \sigma)$$

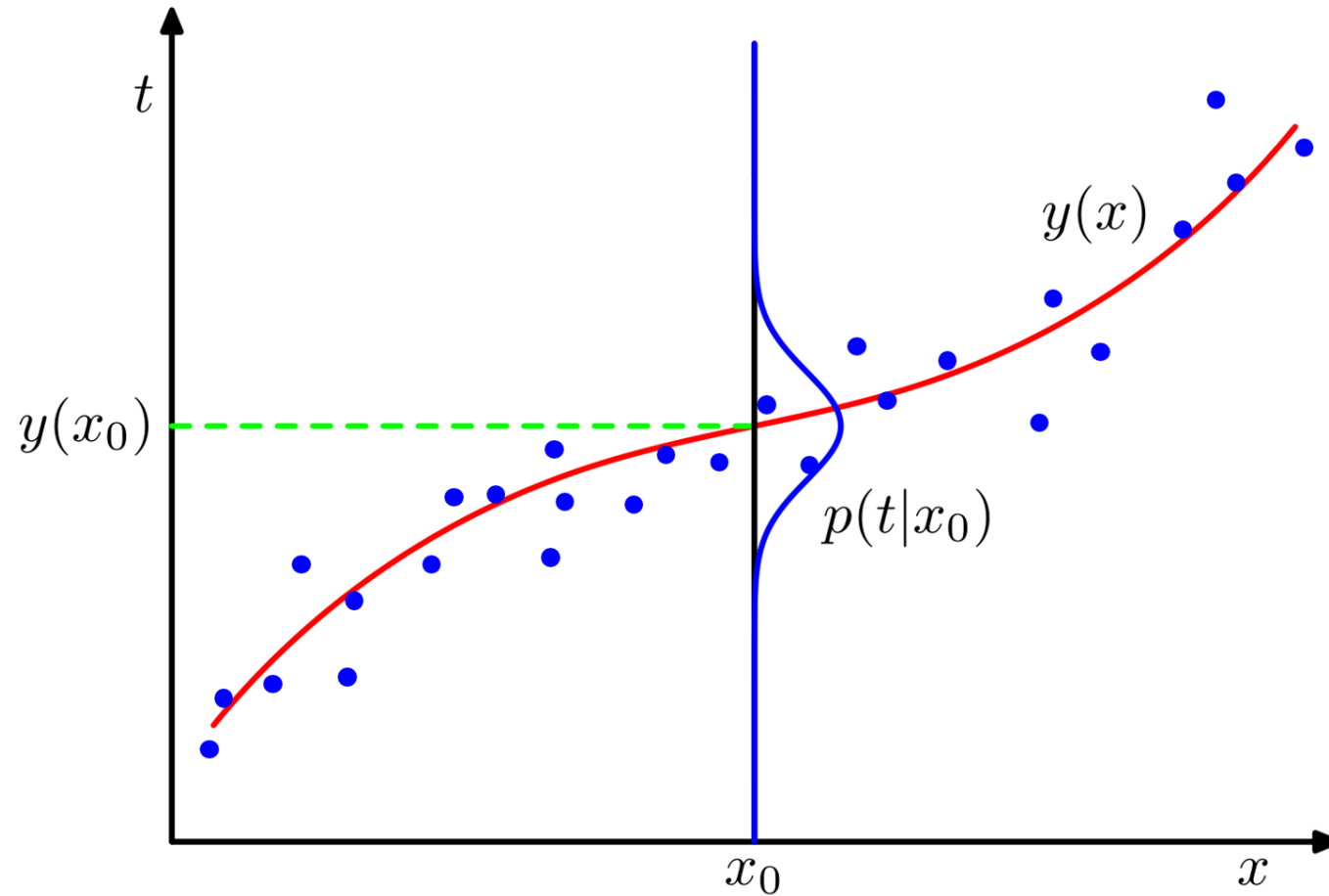
Rethinking Regression Model

cognitive model

statistics

computing

$$\mu_i = \alpha + \beta x_i$$
$$y_i \sim \text{Normal}(\mu_i, \sigma)$$



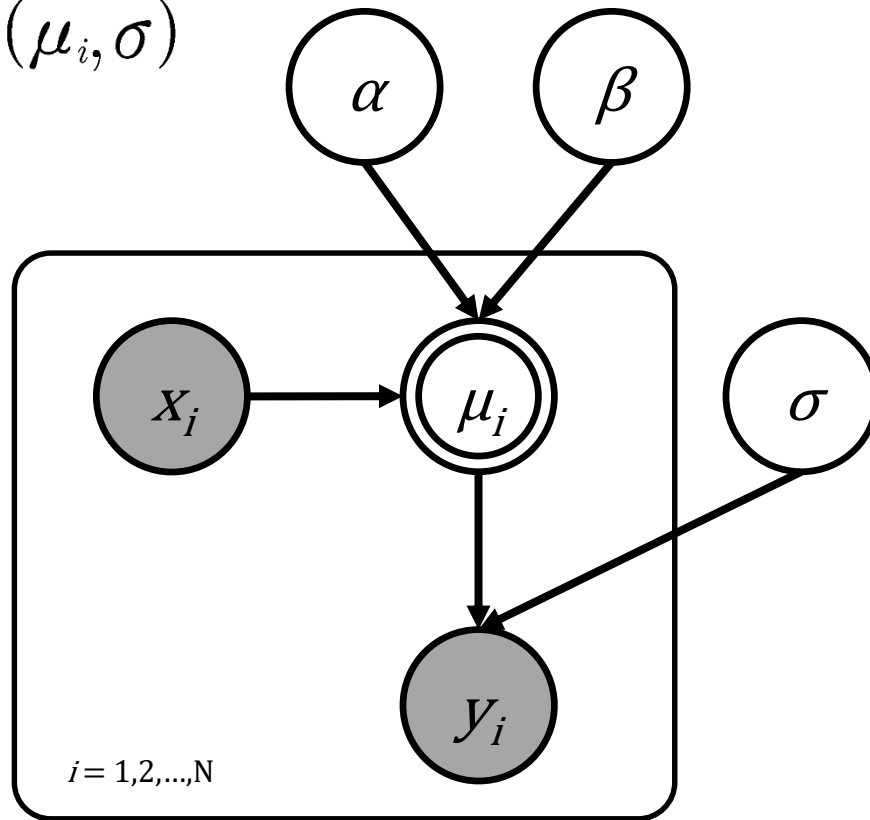
Rethinking Regression Model

cognitive model

statistics

computing

$$\mu_i = \alpha + \beta x_i$$
$$y_i \sim \text{Normal}(\mu_i, \sigma)$$



```
model {  
  vector[N] mu;  
  for (i in 1:N) {  
    mu[i] = alpha + beta * weight[i];  
    height[i] ~ normal(mu[i], sigma);  
  }  
}
```

```
model {  
  vector[N] mu;  
  mu = alpha + beta * weight;  
  height ~ normal(mu, sigma);  
}
```

```
model {  
  height ~ normal(alpha + beta * weight, sigma);  
}
```

Thinking about Priors?

cognitive model

statistics

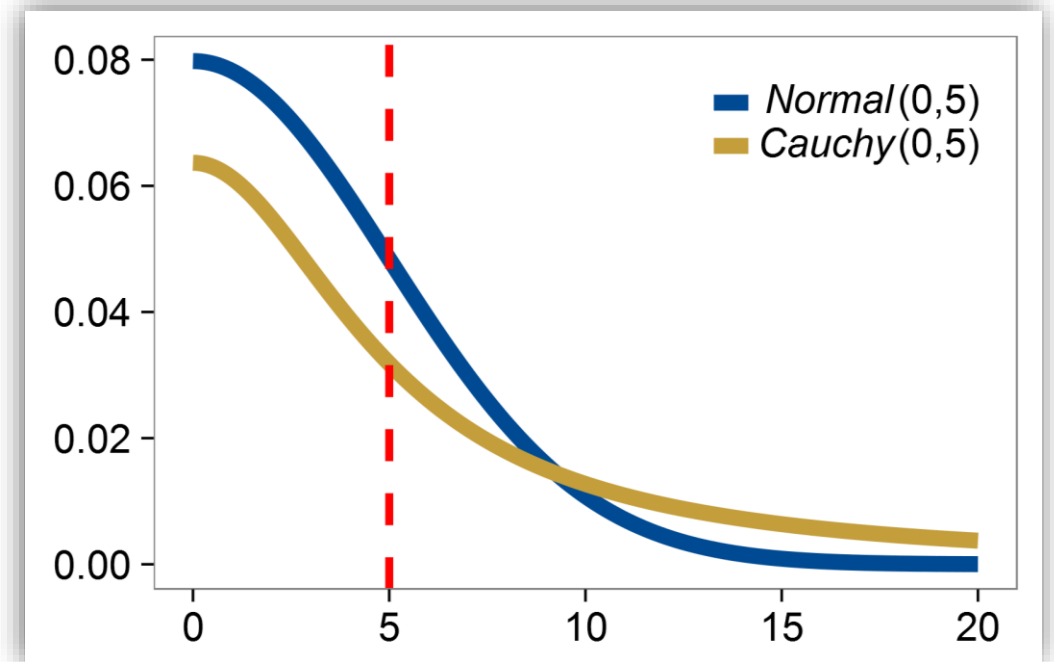
computing

$$\alpha \sim \text{Normal}(170, 100) \quad \beta \sim \text{Normal}(0, 20)$$

$$\overline{\text{height}} = \alpha + \beta * \text{weight}$$

$$\sigma \sim \text{halfCauchy}(0, 20)$$

$$\text{height} \sim \text{Normal}(\overline{\text{height}}, \sigma)$$



Exercise VIII

cognitive model

statistics

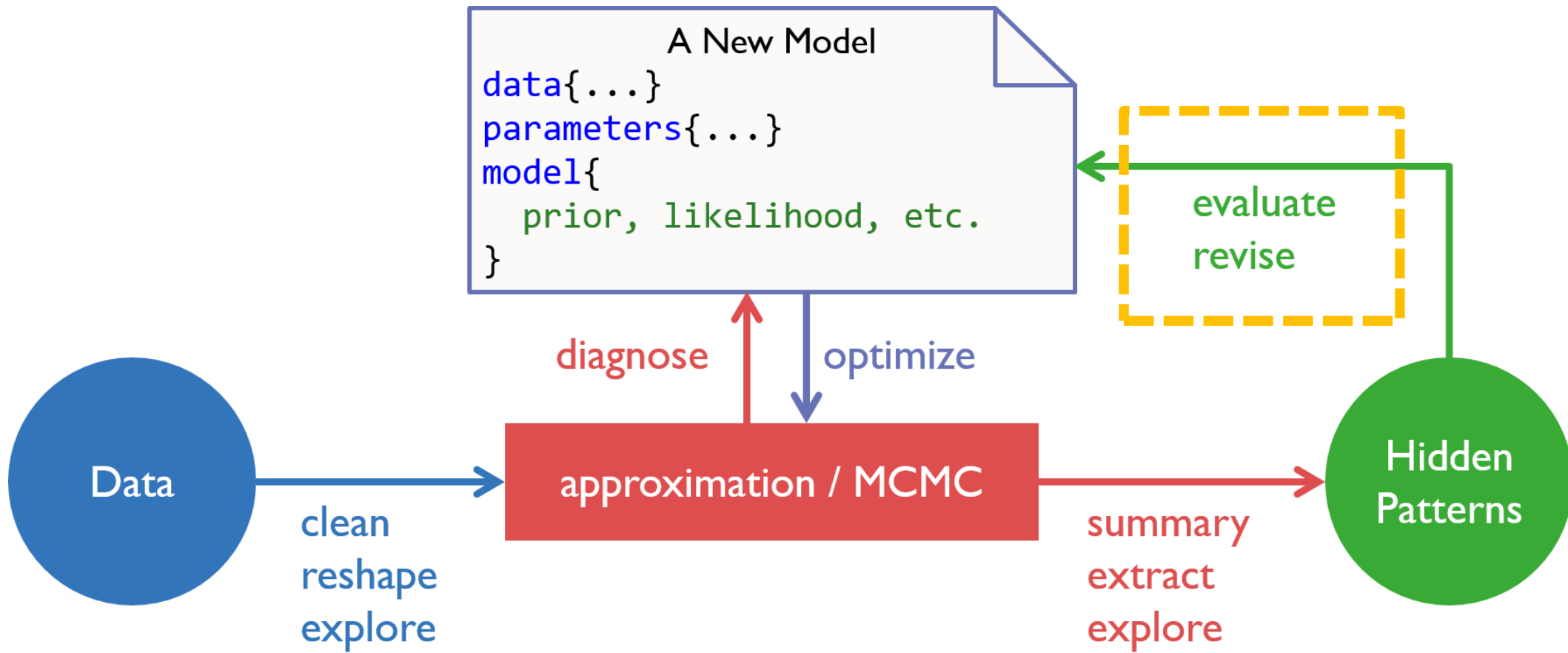
computing

```
.../04.regression_height/_scripts/regression_height_main.R
```

TASK: estimate the model and produce the results

Inference for Stan model: regression_height_model.
4 chains, each with iter=2000; warmup=1000; thin=1;
post-warmup draws per chain=1000, total post-warmup draws=4000.

	mean	se_mean	sd	2.5%	25%	50%	75%	97.5%	n_eff	Rhat
alpha	113.97	0.06	1.86	110.27	112.76	113.93	115.20	117.66	934	1
beta	0.90	0.00	0.04	0.82	0.88	0.90	0.93	0.99	922	1
sigma	5.11	0.01	0.19	4.74	4.97	5.10	5.24	5.50	1437	1
lp__	-747.61	0.04	1.23	-750.80	-748.15	-747.28	-746.72	-746.24	993	1

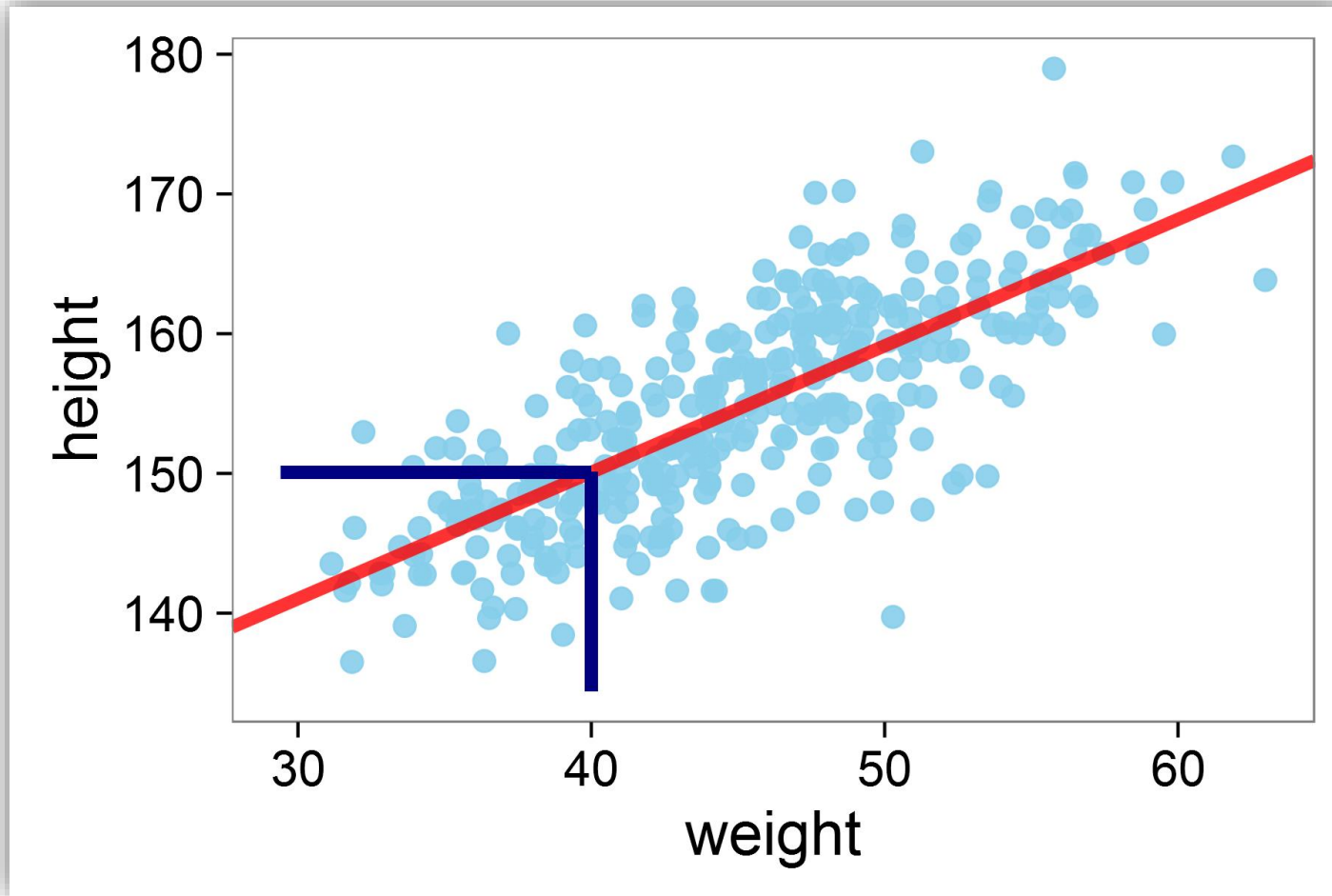


What does the Model Predict?

cognitive model

statistics

computing



$$p(y_{rep} | y) = \int p(y_{rep} | \theta) p(\theta | y) d(\theta)$$

Posterior Predictive Check (PPC)

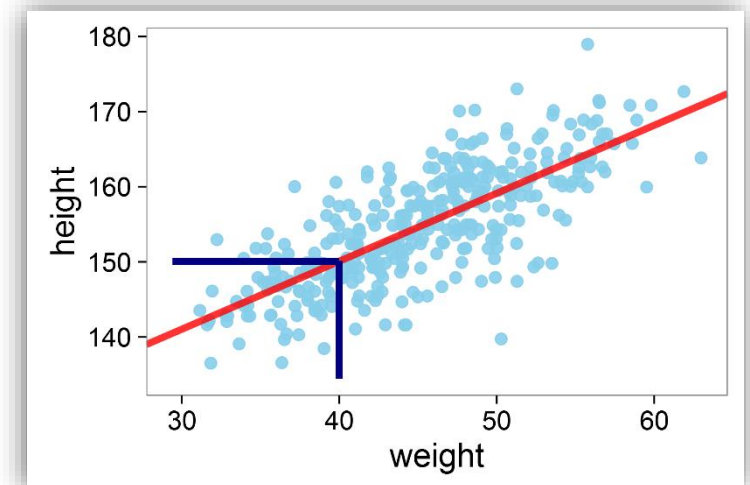
cognitive model

statistics

computing

```
generated quantities {  
  vector[N] height_bar;  
  for (n in 1:N) {  
    height_bar[n] = normal_rng(alpha + beta * weight[n], sigma);  
  }  
}
```

the generated quantities
block runs only AFTER the
sampling, and the time it costs
can be essentially ignored!



Posterior Predictive Check (PPC)

cognitive model

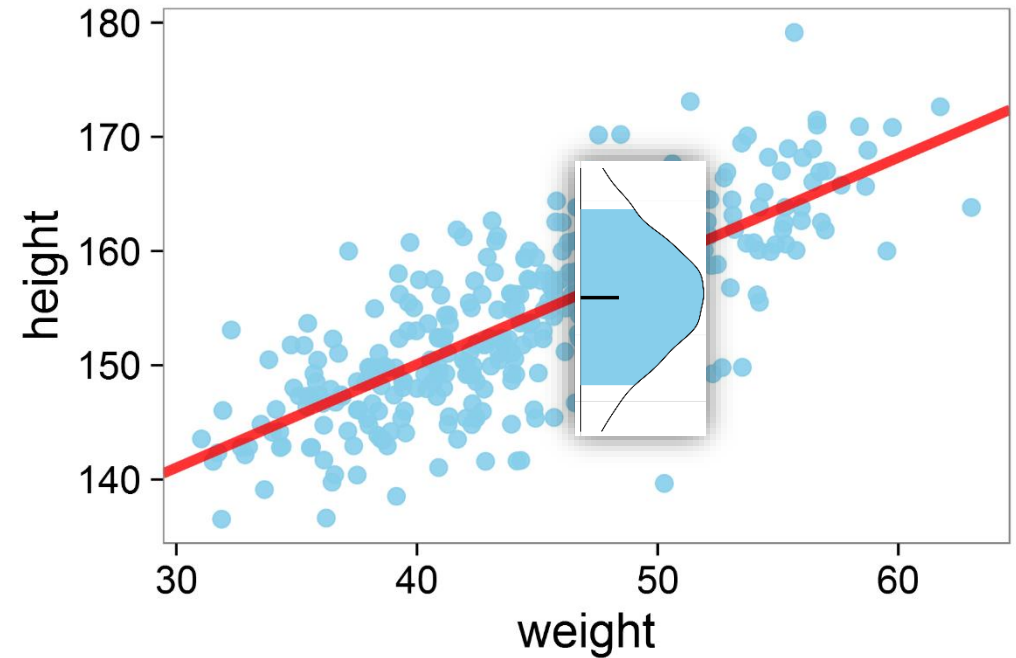
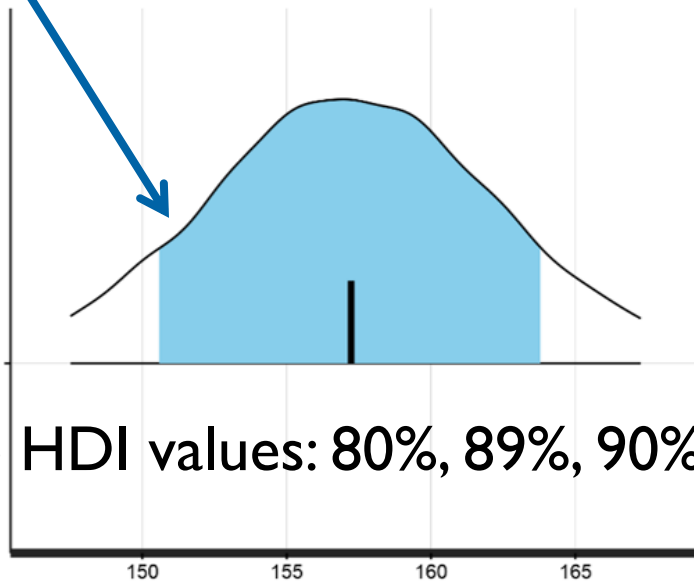
statistics

computing

Highest density interval
(HDI)

`dens(height_bar | x=47.8)`

possible HDI values: 80%, 89%, 90%, 95%



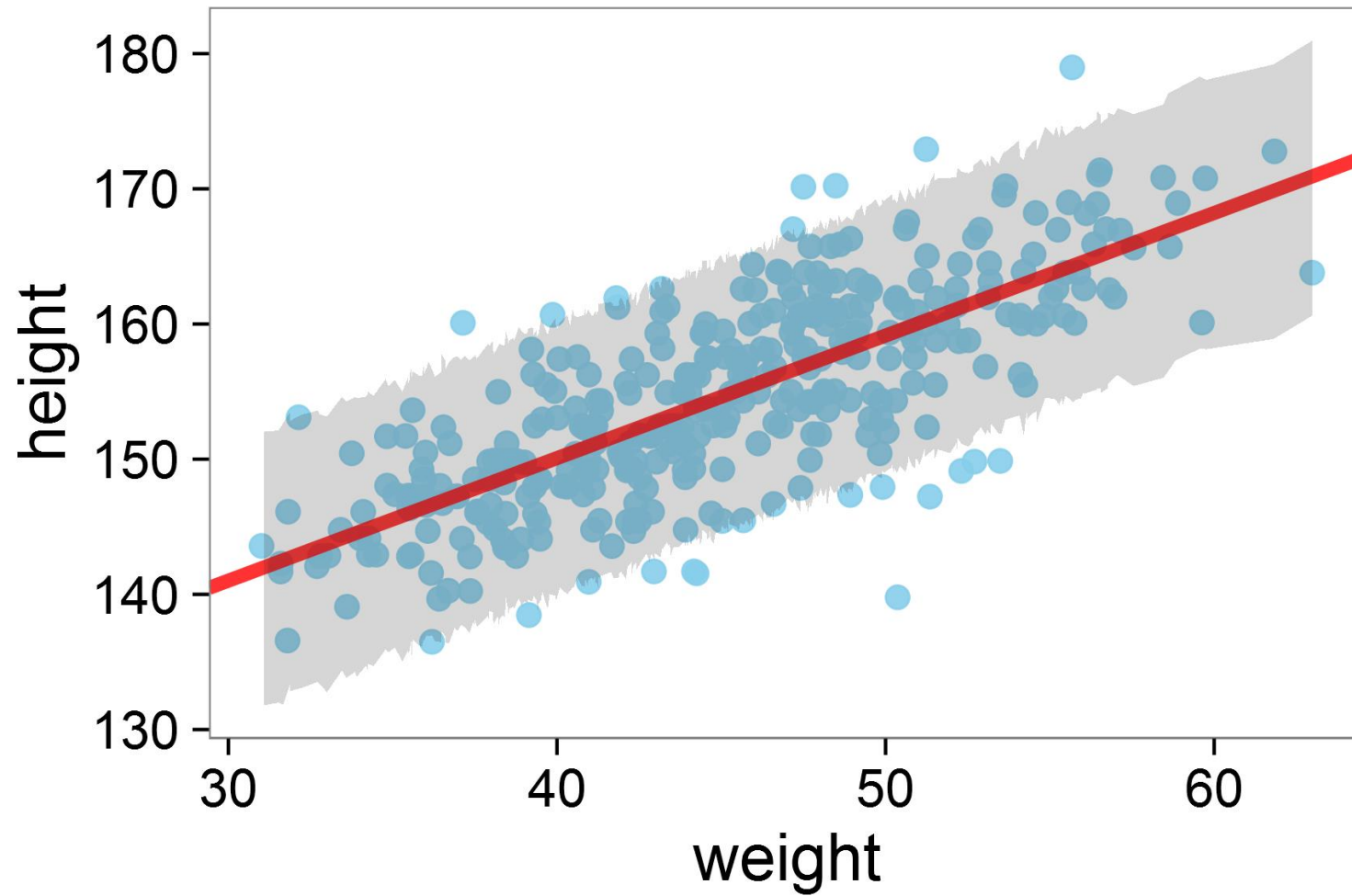
```
height_bar <- extract(fit_reg_ppc, pars = 'height_bar',  
                      permuted = FALSE)$height_bar  
height_HDI <- apply(height_bar, 2, HDIoofMCMC)
```

Posterior Predictive Check (PPC)

cognitive model

statistics

computing



Exercise IX

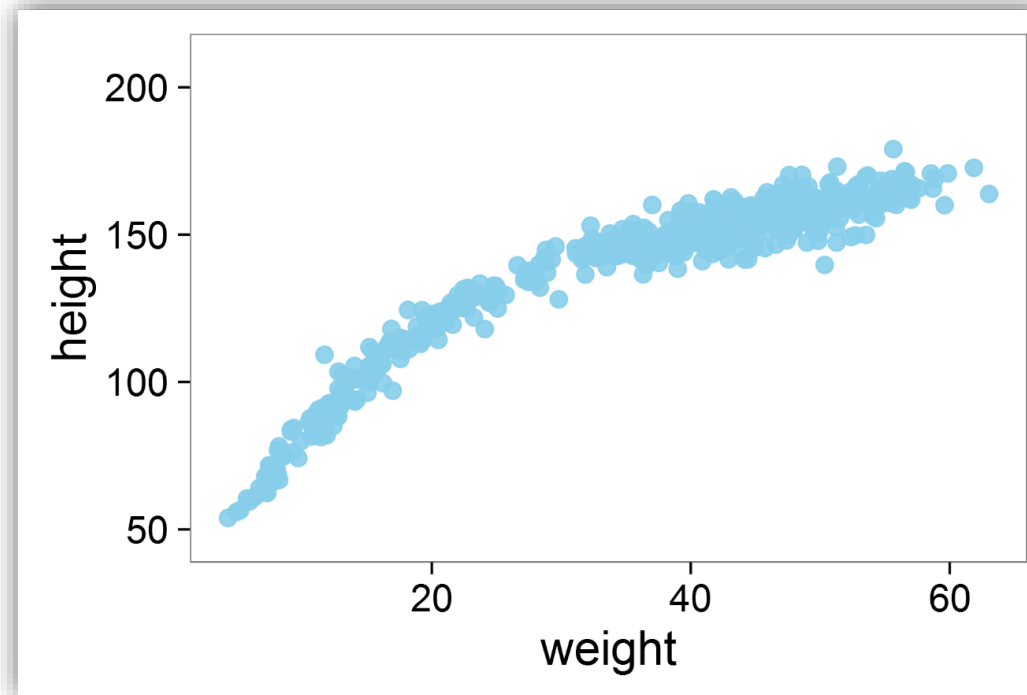
cognitive model

statistics

computing

```
.../05.regression_height_poly/_scripts  
/regression_height_poly_main.R
```

- TASK: (1) Complete “regression_height_poly2_model.stan”
(2) produce PPC plot for both 1st order and 2nd order polynomial fit



Exercise IX – Tips

cognitive model

statistics

computing

```
> source('_scripts/regression_height_poly_main.R')
```

```
> out1 <- reg_poly(poly_order = 1)
```

$$\overline{\text{height}} = \alpha + \beta_1 * \text{weight} + \beta_2 * \text{weight}^2$$

$$\text{height} \sim \text{Normal}(\overline{\text{height}}, \sigma)$$

```
data {  
  int<lower=0> N;  
  vector<lower=0>[N] height;  
  vector<lower=0>[N] weight;  
  vector<lower=0>[N] weight_sq;  
}
```

```
height ~ normal(alpha + beta1 * weight + beta2 * weight_sq, sigma);
```

Exercise IX – output2

cognitive model

statistics

computing

