

The background of the slide features a stylized globe on the left side, partially obscured by a dense pattern of binary code (0s and 1s) that recedes into the distance, creating a sense of depth. The overall color palette is a mix of light blue, teal, and white.

Foundations of Data Science:

Logistic regression -

Multiple logistic regression

Principle of multiple logistic regression

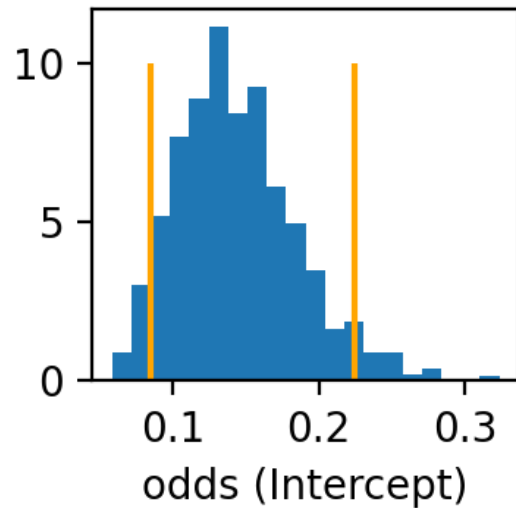
Indep. variables $x^{(1)}$ — Age
 $x^{(2)}$ — Employment
 \vdots \vdots

$$P(Y=1 | x^{(1)}, x^{(2)}, \dots) = f(\beta_0 + \beta_1 x^{(1)} + \beta_2 x^{(2)} + \dots)$$

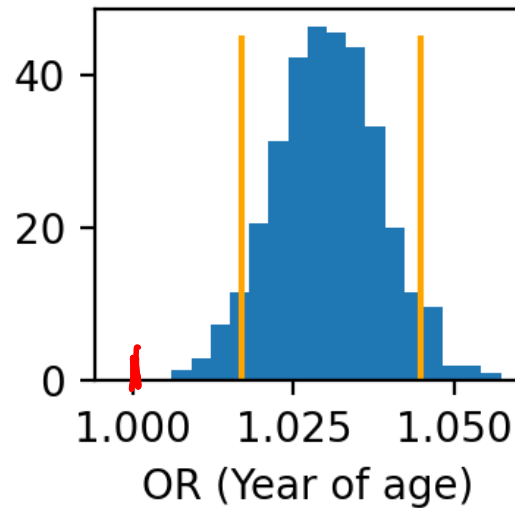
Multiple logistic regression applied to credit e.g.

	Variable	Coefficient	Odds or OR
$\hat{\beta}_0$	Intercept	-1.969	0.140
$\hat{\beta}_1$	Age	0.029	1.030
$\hat{\beta}_2$	Employed	1.881	6.562

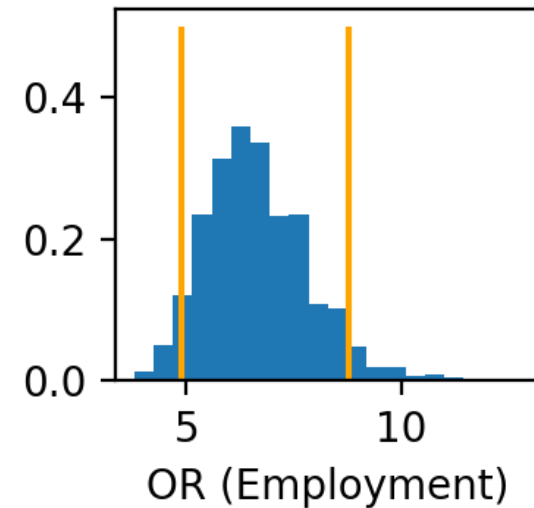
Bootstrap confidence intervals



$$e^{\hat{\beta}_0}$$



$$e^{\hat{\beta}_1}$$



$$e^{\hat{\beta}_2}$$

$$R = 1000$$

$$n = 653$$

H_0 : Age does not affect credit approval

H_a : Age does affect credit approval in some way.

$$H_0 \Rightarrow e^{\hat{\beta}_1} = 1$$

$$p \leq 0.001$$