Basic concepts of Statistics: Inference

Livio Finos

18 October 2018

Contents

Outline **2**
Outline 2
Before we start (in R) 2
The Age vs Reaction Time Dataset 2

Measures of Dependence and the Simple linear model **3**
Measuring the dependence 3
Covariance and Variance 3
Correlation 3
Linear Trend, the least squares method 4
Interpretation of the coefficients 4

Permutation approach to Hypothesis Testing **5**
Some remarks 5
Permutation tests - in a nutshell 5
All potential datasets 5
Random permutations 6
How likely WAS $\hat{\beta}_1^{obs}$? 6
Calculation of the p-value 6
Interpretation 7
We make it in `flip` 7
Composite alternatives (bilateral) 7
Some remarks 8

Parametric Linear Model **8**
From permutation tests (nonparametric) to parametric tests 8
The (simple) linear model 8
Hypothesis testing 9
Power is nothing without control 9
Terminology 9
Properties 10
Confidence intervals 10

A simulation to better understand **10**
A single fictitious dataset 10
H_0 is true (Type I error control) 10
Many datasets 11
H_1 is true (Power evaluation) 12
Many datasets 12
Homework 1: Effect of measure quality (noise) 13
Solution 13
Homework 2: Effect of sample size 14
Solution 14
Homework 3: Confidence intervals 15
Solution: Sigma (noise) 15

Solution: sample size 16

Outline

Outline

Measuring the dependence among variables:

- Covariance and Correlation

- (simple) Linear model

Inference:

- What is about
- Hypothesis testing
- Confidence intervals
- Simulation

Before we start (in R)

```
#clean the memory
rm (list=ls ())

# We customize the output of our graphs a little bit
par.old=par ()
par (cex.main=1.5, lwd=2, col="darkgrey", pch=20, cex=3)
# par (par.old)
palette (c ("#FF0000", "#00A08A", "#FFCC00", "#445577", "#45abff"))

# customize the output of knitr
knitr :: opts_chunk$set (fig.align="center")#, fig.width=6, fig.height=6)
```

The Age vs Reaction Time Dataset

The reaction time of these subjects was tested by having them grab a meter stick after it was released by the tester. The number of centimeters that the meter stick dropped before being caught is a direct measure of the person's response time.

The values of `Age` are in years. The `Gender` is coded as `F` for female and `M` for male. The values of `Reaction.Time` are in centimeters.

(data are fictitious)

To read the data

```
data(reaction,package = "flip")
# or download it from: https://github.com/livioivil/flip/tree/master/data
# str (reaction)
```

We plot the data

```r
plot(x=reaction$Age,y=reaction$Reaction.Time,pch=20,col=2,cex=2)
```

Measures of Dependence and the Simple linear model

Measuring the dependence

we define:

- $X = Age$

- $Y = Reaction.Time$

We review some famous index to measure the (linear) dependence among two variables

Covariance and Variance

Covariance between X and Y:

$\sigma_{xy} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{n}$

- values between $-\infty$ and ∞

- $\sigma_{xy} \approx 0$: there is no dependency between X and Y

- $\sigma_{xy} >> (<<) 0$: there is a strong positive (negative) dependency between X and Y

Variance of X (= covariance between X and X):

$\sigma_{xx} = \sigma_x^2 = \frac{\sum_{i=1}^{n}(x_i-\bar{x})^2}{n}$

Standard Deviation of X:

$\sigma_{xx} = \sqrt{\sigma_{xx}} = \sigma_x$

Correlation

With the Covariance it is difficult to understand when the relationship between X and Y is strong / weak. We note that

$-\sigma_x\sigma_y \leq \sigma_{xy} \leq \sigma_x\sigma_y$ is quivalent to $-1 \leq \frac{\sigma_{xy}}{\sigma_x\sigma_y} \leq 1$

Correlation between X and Y:

$\rho_{xy} = \frac{\sigma xy}{\sigma_x\sigma_y} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i-\bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i-\bar{y})^2}}$

- values between -1 and 1
- $\rho_{xy} \approx 0$: there is no dependency between X and Y
- $\rho_{xy} \approx 1(-1)$: there is a strong positive (negative) dependency between X and Y

Linear Trend, the least squares method

We describe the relationship between
`Reaction.Time` and `Age` with a straight line.

$Reaction.Time \approx \beta_0 + \beta_1 Age$
$Y = \beta_0 + \beta_1 X$

Let's draw a line 'in the middle' of the data.

The **least-squares estimator**

We look for the one that passes more 'in the middle', the one that minimizes the sum of the squares of the residues:

$\hat{\beta}_0$ and $\hat{\beta}_1$ such that
$\sum_{i=1}^{n}(y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i))^2$ is minimum.

Estimates:

- Angular coefficient: $\hat{\beta}_1 = \frac{\sigma_{xy}}{\sigma_{xx}} = \rho_{xy}\frac{\sigma_y}{\sigma_x} = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2} = 0.2064719$
- Intercept: $\hat{\beta}_0 = \bar{y} - \hat{\beta}_1\bar{x} = 10.3013483$
- Response (estimated y): $\hat{y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_i$
- Residuals (from the estimated response): $y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i) = y_i - \hat{y}_i$

and therefore the least squares are the sum of the squared residuals: $\sum_{i=1}^{n}(y_i - \hat{\beta}_0 + \hat{\beta}_1 x_i)^2 = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$

A graphical representation:

```
model=lm(Reaction.Time~Age,data=reaction)
coefficients(model)
```

```
## (Intercept)         Age
##  10.3013483   0.2064719
```

```
plot(reaction$Age,reaction$Reaction.Time,pch=20,col=2,cex=1)
coeff=round(coefficients(model),1)
title(paste("Y=",coeff[1],"+",coeff[2],"*X"))
abline(model,col=1)
```

Interpretation of the coefficients

- β_0 indicates the value of y when $x = 0$ (where the line intersects the ordinate axis).
- β_1 indicates how much y grows as a unit of x grows
  - If $\beta_1 = 0$ there is no relation between x and y.$Y$is constant (horizontal), knowing$x$does not change the estimate of y
  - If $\beta_1 > (<)0$ the relation between x and y is positive (negative). When X passes from x a $x+1$ the estimate of Y changes from $\hat{y}$ to $\hat{y} + \hat{\beta}_1$

Permutation approach to Hypothesis Testing

Some remarks

Let's note that all the measures above does not make any assumptions on the random process that generate them.

Let's assume that Y - and possibly X - is not fix, while it is generated by a random variable.

The question: **Is there a relationship between Y and X?**

We estimated $\hat{\beta}_1 = 0.2064719$

but the **true value** β_1 is really different from 0 (i.e. no relationship)?
Otherwise, is the distance to 0 is due to the random sampling?

- **Null Hypothesis** $H_0: \ \beta_1 = 0$ (the **true** β_1, not its estimate $\hat{\beta}_1$!). There is no relationship between X and Y.
- **Alternative Hypothesis** $H_1: \ \beta_1 > 0$ The relationship is positive.

Other possible specifications of $H_1: \ \beta_1 < 0$ and, more commonly, $H_1: \ \beta_1 \neq 0$.

Permutation tests - in a nutshell

As a toy example, let use a sub-set of the data:

```
##   Age Gender Reaction.Time
## 3  30      M         11.62
## 4  60      F         22.27
## 5  80      M         26.48
```

- *If* H_0 is true: there is no linear relationship between X and Y
- Therefore, the trend observed on the data is due to chance.
- Any other match of x_i and y_i was equally likely to occur
- I can generate the datasets of other hypothetical experiments by exchanging the order of the observations in Y.
- How many equally likely datasets could I get with$X$ and Yobserved? $3*2*1 = 3! = 6$ possible datasets.

Remark: Here we only assume that y is a random variable. The only assumption here is the exchangeability of the observations: the joint density $f(y_1, \ldots, y_n)$ does not change when the ordering of $y_1, \ldots, y_n$ is changed.

All potential datasets

Random permutations

In our data set, if we apply the same principle…

How many permutations of the vector $y_1, \ldots, y_n$ are possible? $n! = 3628800$.

big, perhaps not too big … but what happen with, for example, $n = 20$? We got $20! = 2.432902e + 18$. This is too big, definitely!

We calculate a smaller (but sufficiently large) B of random permutations.

here some example

`Age` vs a permutations of `Reaction.Time`

We repeat 10ˆ{4} times and we look at the histogram of the $\hat{\beta}_1$

```r
# beta_1 estimated on the observed data:
beta1=coefficients(lm(Reaction.Time~Age,data=reaction))[2]

# function that permutes the y values and calculates the coeff beta_1
my.beta.perm <- function(Y,X){
  model=lm(sample(Y)~X)
  coefficients(model)[2]
}

#replicate it B-1 times
beta.perm= replicate(B,my.beta.perm(reaction$Reaction.Time, reaction$Age ))
```

How likely WAS $\hat{\beta}_1^{obs}$?

(before the experiment!)

How likely was it to get a $\leq \hat{\beta}_1^{obs}$ value among the many possible values of $\hat{\beta}_1^{*b}$ (obtained by permuting data)?

Remarks:

- $\hat{\beta}_1^{*b} < \hat{\beta}_1^{obs}$ (closer to 0): less evidence against H_1 than $\hat{\beta}_1^{obs}$
- $\hat{\beta}_1^{*b} \geq \hat{\beta}_1^{obs}$: equal or more evidence towards H_1 than $\hat{\beta}_1^{obs}$

Calculation of the p-value

Over B=10ˆ{4} permutations we got 9822 times a $\hat{\beta}_1^{*b} \leq \hat{\beta}_1^{obs}$.

The p-value (significance) is $p = \frac{\#(\hat{\beta}_1^{*b} \geq \hat{\beta}_1^{obs})}{B+1} = 0.018$

Interpretation

The probability of $p = P(\hat{\beta}_1^* \leq \hat{\beta}_1 = 0.206\ |H_0)$ is equal to $p = 0.018$, i.e. very small.
So, it was unlikely to get a value like this **IF H_0 is true**.

Neyman-Pearson's approach has made common the use of a significance threshold for example $\alpha = .05$ (or $= .01$). When $p \leq \alpha$ rejects the hypothesis that there is no relationship between X and Y (H_0). If so, we are inclined to think that H_1 is true (there is a positive relationship).

- Type I error: False Positive
  the true hypo is H_0 (null correlation), BUT we accept H_1 (correlation is positive)
- Type II error: False Negative
  the true hypo is H_1 (positive correlation), BUT we do not reject H_0 (null correlation)

Type I error control

We want to guarantee not to get false relationships (a few false positives), better to be conservative. To make this, we want to bound the probability to make a false discovery:

$P(p - value \leq \alpha | H_0) \leq \alpha$

We built a machinery that in the long run (many replicates of the experiment) finds false correlations with probability α (e.g. $0.05 = 5\%$).

We make it in `flip`

```r
library(flip)

(res=flip(Reaction.Time~Age,data=reaction,
          tail=1,perms = 10000))
```

```
##
##
##                 Test  Stat tail p-value
## Reaction.Time    t 2.633    >  0.0158
```

```r
## compare also with
# flip(Reaction.Time~Age,data=reaction,tail=1,statTest = "cor")
# flip(Reaction.Time~Age,data=reaction,tail=1,statTest = "coeff")
```

```r
plot(res)
```

Composite alternatives (bilateral)

The hypothesis $H_1:\ \beta_1 > 0$ (the relation is positive) must be justified with a priori knowledge.

More frequently, the Alternative hypothesis is appropriate: $H_1:\ \beta_1 \neq 0$ (there is a relationship, I do not assume the direction)

I consider anomalous coefficients estimated as very small but also very large ('far from 0'). The p-value is $p = \frac{\#(|\hat{\beta}_1^{*b}| \geq |\hat{\beta}_1^{obs}|)}{B+1} = 0.0345$

In `flip`:

```r
library(flip)
(res=flip(Reaction.Time~Age,data=reaction,tail=0,perms=5000))
```

```
##

##
##                Test  Stat tail p-value
## Reaction.Time    t 2.633   ><  0.0398
```

```r
plot(res)
```

Some remarks

- Do not be confused with bootstrap methods. The former are extractions without reintegration, the latter with. The former have almost optimal properties and have (almost always) an exact control of the first type errors.
- A general approach and are applicable in many contexts. Very few assumptions.
- Some dedicated R packages:
  - flip (the development version is on github)
  - coin
  - permuco
- They are of limited applicability when there are many variables involved.

Parametric Linear Model

From permutation tests (nonparametric) to parametric tests

We can see that the histogram of the statistical tests (calculated on the permuted data) is well described by a **Gaussian** (normal) curve.

```r
hist(beta.perm,50,probability=TRUE,col=2)
curve(dnorm(x,mean(beta.perm),sd(beta.perm)),add=TRUE,col=1,lwd=3)
points(beta1,0,lwd=3,col=1)
```

The (simple) linear model

We assume that the observed values are distributed around true values $\beta_0 + \beta_1 X$ according to a Gaussian law:

$Y = \text{linear part} + \text{normal error}$

$Y = \beta_0 + \beta_1 X + \varepsilon$

Assumptions of the linear model

- the $\boldsymbol{y_i = \beta_0 + \beta_1 x_i + \varepsilon_i}$ the relationship between X and the true (mean) Y is linear.

- the **observations** are **independent** each others (knowing the value of the $y_i$observation does not help me to predict the value of y_{i+1}). The random part is ε_i, these are the independent terms.
- $\boldsymbol{\varepsilon_i \sim N(0, \sigma^2),\ \forall i = 1, \dots, n}$ errors have normal distribution with zero mean and common variance (homoschedasticity: same variance).

Hypothesis testing

If these assumptions are true,

$\hat{\beta_1} \sim N(\beta_1, \sigma^2 / \sum(x_i - \bar{x})^2)$

We calculate the test statistic:

$t = \frac{\hat{\beta_1}}{std.dev\ \hat{\beta_1}} = \frac{\hat{\beta_1}}{\sqrt{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 / \sum(x_i - \bar{x})^2 / (n-2)}}$

If $H_0 : \beta_1 = 0$, $t \sim t(n-2)$ is true

On `reaction` data and $H_1 : \beta_1 \neq 0$ (bilateral alternative)

```
model=lm (Reaction.Time ~ Age, data=reaction)
summary(model) $ coefficients
```

```
##               Estimate Std. Error  t value   Pr(>|t|)
## (Intercept) 10.3013483 4.04406774 2.547274 0.03431997
## Age          0.2064719 0.07841111 2.633197 0.03002876
```

Similar result, but much more assumptions!

Power is nothing without control

We don't know if the data are genareted under H_0 or under H_1.

But we have a tool (the test) that

- if the data are generated **under** H_0: it suggests (wrong) H_1 (i.e. $p \leq \alpha$, type I error, false positive) with probability α. E.g. $\alpha = .05$, low probability.
- if the data are generated **under** H_1: it suggests (correct) H_1 (i.e. true positive) with probability larger than α.
  This is the Power of a test. The Power is unknown, but we hope it is as high as possible.

Terminology

- Probability of **Type I error** (Probability of **False Positive**, α): the probability to find a relationship when it does not exist (true H_0, the test judges H_1).

- Probability of **Type II error** (Probability of **False Negative**): the probability NOT to find a relationship when it does exist (true H_1, the test judges H_0).

- **Specificity**: the probability NOT to find a relationship when it does NOT exist (true H_0, the test judges H_0). it is equal to 1 - Type I error.

- **Power** (**Sensitivity**): the probability to NOT find a relationship when it does exist (true H_1, the test judges H_1). it is equal to 1 - Type II error.

Properties

If the parametric assumptions are valid, the test guarantes

- the control of the type I error at the α level,

- the maximum power (minimum error of type II β) among all the possible tests,

- asymptotic consistency (if they are under H_1 rejection always for sufficiently large n).

The permutation tests usually have slightly less power and converge to the corresponding parametric tests, IF they exist.

Confidence intervals

The parametric approach also allows us to calculate confidence intervals

```
confint(model)
```

```
##                  2.5 %     97.5 %
## (Intercept) 0.97571138 19.6269853
## Age         0.02565557  0.3872883
```

In linear model:

$$C.I. = [\hat{\beta}_1 - t_{1-\alpha/2}\hat{\sigma}/\sqrt{n}, \hat{\beta}_1 + t_{1-\alpha/2}\hat{\sigma}/\sqrt{n}]$$

($t_{1-\alpha/2}$ is the threshold given by a t-distribution with CDF equal to $1 - \alpha/2$ and d.f. $n - 1$)

Confidence intervals are constructed in such a way that in the long run they include the true value β_1 with probability $1 - \alpha$ (e.g. 95 %).

Once the data has been collected, the Conf Int is computed.
It will includes or not the true value β_1.
We only have the certificate of quality of out test, Conf Int in the 95% of the (previous) cases was wrong 95% of the times.

Confidence intervals and hypothesis testing are closely related: if a confidence interval at level $1 - \alpha$ does not include the 0, the p-value that tests $H_0: \ \beta_1 = 0$ will be $p < \alpha$.

A simulation to better understand

A single fictitious dataset

We generate a fictitious datasets and see how the tests and confidence intervals behave.

- use the observed values for Age in the original dataset

- randomly generate values for $Reaction.Time$.
- get the statistics (p-values, confidence interval)

H_0 is true (Type I error control)

There is no relationship between Age and $Reaction.Time$.

then: $Reaction.Time = \beta_0 + 0Age + \varepsilon$

ε can be assumed to be normal $N(0, \sigma^2)$.

How to set β_0 and σ^2?
As reasonable values, we can use mean and variance calculated on the sample:

```
m=mean(reaction$Reaction.Time)
s2=var(reaction$Reaction.Time)
s=sqrt(s2)
n=length(reaction$Age)
```

- $\beta_0 = 20.212$
- $\sigma^2 = 36.3187511$

```
# generate random Reaction.Time
Reaction.Time=rnorm(n,m,s) # equivalent to m+rnorm(n,0,s)
#and fit the model
mod=lm(Reaction.Time~reaction$Age)
summary(mod)$coefficients
```

```
##                Estimate Std. Error   t value    Pr(>|t|)
## (Intercept)  25.9745219  5.5037753  4.719401 0.001503326
## reaction$Age -0.1076552  0.1067136 -1.008823 0.342594709
```

```
confint(mod)
```

```
##                    2.5 %     97.5 %
## (Intercept)  13.2827934 38.6662505
## reaction$Age -0.3537373  0.1384269
```

Many datasets

Now we generate many (e.g. 1000) datasets and we store the p-values.

```
sim <- function(Age,n,m,s){
  Reaction.Time=rnorm(n,m,s)
  #get the p-value from the output
  summary(lm(Reaction.Time~Age))$coefficients["Age","Pr(>|t|)"]
}

p.sim=replicate(1000,sim(reaction$Age,n,m,s))
```

- What do I expect the distribution of these p-values to be?
- If I plot a histogram, what do I expect?
- What will be the proportion of p-values ≤ 0.05?

```
hist(p.sim)
```

```
#how many p<.05
sum(p.sim<.05)
```

```
## [1] 51
```

```
#proportion of p<.05
mean(p.sim<.05)
```

```
## [1] 0.051
```

Now the (empirical) cumulative distribution

```
plot(ecdf(p.sim),xlab="alpha",ylab="Empirical Type I error", col=1,main="Type I error as a function of
abline(0,1,col=2)
```

For each value of the abscissa α we see the empirical estimate of the type I error.

For any given α, the estimated proportion is around α (and would converge to α if we increases the number of replications)

Very good!! :)

H_1 is true (Power evaluation)

There is a relationship between *Age* and *Reaction.Time*. I can use use the linear normal model with parameters - just an example - calculated on the sample.

```
modelF=lm(Reaction.Time~Age,data=reaction)
coefficients(modelF)
```

```
## (Intercept)         Age
##  10.3013483   0.2064719
```

```
beta0=coefficients(modelF)[1]
beta1=coefficients(modelF)[2]
```

```
y_linear_part=beta0+beta1*reaction$Age
s=sd(residuals(modelF))
n=length(reaction$Age)

# generate random Reaction.Time
Reaction.Time=y_linear_part+rnorm(n,0,s)
#fit the model (estimate the parameters)
summary(lm(Reaction.Time~Age,data=reaction))$coefficients
```

```
##                Estimate Std. Error  t value   Pr(>|t|)
## (Intercept) 10.3013483 4.04406774 2.547274 0.03431997
## Age          0.2064719 0.07841111 2.633197 0.03002876
```

Many datasets

Now we generate many (e.g. 1000) datasets and, for each, we store the p-value.

```
p.sim=replicate(1000,sim(reaction$Age,n,y_linear_part,s))
```

- What do I expect the distribution of these p-values to be?

- If I plot a histogram, what do I expect?
- What will be the proportion of p-values ≤ 0.05?

```
hist(p.sim)
```

```
#how many p<.05
sum(p.sim<.05)
```

```
## [1] 678
```

```
#proportion of p<.05
mean(p.sim<.05)
```

```
## [1] 0.678
```

Now the (empirical) cumulative distribution

```
plot(ecdf(p.sim),xlab="alpha",ylab="Empirical Power", col=1,main="Power as a function of alpha",lwd=3,a
abline(0,1,col=2)
```

For each value of the abscissa α we see the empirical estimate of the Power.

For any given α, the estimated proportion greater than α, the test has power!

Very good!! :)

Homework 1: Effect of measure quality (noise)

1. How do the Type I error varies as a function of the standard deviation of the normal errors?
   Hint: Simulate with different sd (e.g. 2,4,8,16), store the proportion of rejections for $\alpha = .05$, plot the sd vs rejections.

2. Same task, but under H_1 (power study)

Solution

```
sim05 <- function(s,n,m){
  p.sim=replicate(1000,sim(reaction$Age,n,m,s))
  mean(p.sim<.05)
}
```

```
s_values=c(2,4,8,16)
type1=sapply(s_values,sim05,n,m)
plot(s_values,type1,type="l",ylim=c(0,.2),main="Type I error",xlab="Sigma (noise)")
```

```r
power=sapply(s_values,sim05,n,y_linear_part)
plot(s_values,power,type="l",main="Power",xlab="Sigma (noise)")
```

Homework 2: Effect of sample size

1. How does the type I error varies as a function of the sample size? (Same hint of homework 1, but here you sample `Age` `n` times and compute `y_linear_part`)
2. How does the power?

Solution

```r
sim_n <- function(Age,n,betas,s){
  Agex <- sample(Age,n,replace = TRUE)
  while(length(unique(Agex))==1)
    Agex <- sample(Age,n,replace = TRUE)

  y_linear_part=betas[1]+betas[2]*Agex
  Reaction.Time <- rnorm(n,y_linear_part,s)
  #get the p-values from the output
  summary(lm(Reaction.Time~Agex))$coefficients["Agex","Pr(>|t|)"]
}

sim05_n <- function(n,betas,s){
  p.sim=replicate(1000,sim_n(reaction$Age,n,betas,s))
  mean(p.sim<.05)
}
```

H0

```r
betas=c(beta0,0)
ns=c(5,10,20,50)
type1=sapply(ns,sim05_n,betas,s)
plot(ns,type1,type="l",ylim=c(0,.2),main="Type I error",xlab="sample size")
```

H1

```r
betas=c(beta0,beta1)
power=sapply(ns,sim05_n,betas,s)
plot(ns,power,type="l",main="Power",xlab="sample size")
```

Homework 3: Confidence intervals

Make similar evaluations of Homework 1 and 2 for confidence intervals:

- set $\alpha = .05$,
- genate datasets, fit the models,
- Two quantities are of interest here:
  - counts of times the confidence interval contains the TRUE value β_1
  - length of the confidence interval

Solution: Sigma (noise)

```
sim <- function(Age,n,m,s){
  Reaction.Time=rnorm(n,m,s)
  #get the p-value from the output
  CI=confint(lm(Reaction.Time~Age))[2,]
  contained=all(sign(CI)==c(-1,1))
  CI_length=CI[2]-CI[1]
  names(CI_length)=NULL
  c(contained,CI_length)
}

sim05 <- function(s,n,m){
  p.sim=replicate(1000,sim(reaction$Age,n,m,s))
  rowMeans(p.sim)
}
```

H0

```
s_values=c(2,4,8,16)
type1=sapply(s_values,sim05,n,m)
par(mfrow=c(1,2))
plot(s_values,type1[1,],type="l",ylim=c(.8,1),main="Coverage",xlab="Sigma (noise)")
plot(s_values,type1[2,],type="l",main="Length",xlab="Sigma (noise)")
```

H1

```
power=sapply(s_values,sim05,n,y_linear_part)
par(mfrow=c(1,2))
plot(s_values,type1[1,],type="l",ylim=c(.8,1),main="Coverage",xlab="Sigma (noise)")
plot(s_values,type1[2,],type="l",main="Length",xlab="Sigma (noise)")
```

Solution: sample size

```
sim_n <- function(Age,n,betas,s){
  Agex <- sample(Age,n,replace = TRUE)
  while(length(unique(Agex))==1)
    Agex <- sample(Age,n,replace = TRUE)

  y_linear_part=betas[1]+betas[2]*Agex
  Reaction.Time <- rnorm(n,y_linear_part,s)
  CI=confint(lm(Reaction.Time~Agex))[2,]
  contained=all(sign(CI)==c(-1,1))
  CI_length=CI[2]-CI[1]
  names(CI_length)=NULL
  c(contained,CI_length)
}

sim05_n <- function(n,betas,s){
  p.sim=replicate(1000,sim_n(reaction$Age,n,betas,s))
  rowMeans(p.sim)
}
```

H0

```
betas=c(beta0,0)
ns=c(5,10,20,50)
type1=sapply(ns,sim05_n,betas,s)
par(mfrow=c(1,2))
plot(ns,type1[1,],type="l",ylim=c(.8,1),main="Coverage",xlab="n (Sample Size)")
plot(ns,type1[2,],type="l",main="Length",xlab="n (Sample Size)")
```

H1

```
betas=c(beta0,beta1)
power=sapply(ns,sim05_n,betas,s)
par(mfrow=c(1,2))
plot(ns,type1[1,],type="l",ylim=c(.8,1),main="Coverage",xlab="Sigma (noise)")
plot(ns,type1[2,],type="l",main="Length",xlab="Sigma (noise)")
```