# UAS Komputasi Statistik

Rahmawati Annisa Salsadilla - 11180940000005

16/12/2020

## halaman 2

Jelaskan perbedaan antara standard deviasi suatu sampel dan standard error suatu statistik (seperti mean sampel)?

Jawab:

Standar deviasi suatu sampel adalah suatu indeks yang menggambarkan sebaran data terhadap rata-ratanya, sedangkan standard error (mean sampel) adalah indeks yang menggambarkan sebaran rata-rata sampel terhadap rata-rata dari rata-rata keseluruhan kemungkinan sampel (rata-rata populasi).

Jelaskan apa yang dimaksud dengan distribusi sampling ?

Jawab:

Distribusi sampling adalah distribusi peluang teoritis dari ukuran-ukuran statistik, misalnya rata-rata,varians dan proporsi yang mungkin muncul dari sample-sampel.

## halaman 3 simulasi monte carlo

hitung berapa P(T>2)?

```
N <- 10000
set.seed(005)
x <- rexp(n=N, rate = 1)
y <- rexp(n=N, rate = 1/2)
z <- rexp(n=N, rate = 1/3)
for (i in 1:N) {
  T <- sqrt(x*y*z)
  peluang <- (T>2)
}
nilai_peluang <- sum(peluang)/N
nilai_peluang
```

```
## [1] 0.3005
```

Estimasi berapa variansi dari T?

```
N <- 10000
set.seed(005)
x <- rexp(n=N, rate = 1)
y <- rexp(n=N, rate = 1/2)
z <- rexp(n=N, rate = 1/3)
for (i in 1:N) {
  T <- sqrt(x*y*z)
  peluang <- (T>2)
```

```
}
var(T)
```

```
## [1] 3.135771
```

## halaman 4

dua buah dadu dilempar secara bersamaan sampai muncul sisi angka yang sama. Misalkan T adalah banyaknya pelemparan dadu. Dengan menggunakan simulasi monte carlo estimasi berapa E[T]? lakukan simulasi sebanyak 10 ribu kali.

```
set.seed(005)
N <- 10000
n <- 10
T <- rep(NA,N)
for (i in 1:N) {
  pelemparan_dadu <- replicate(2, sample(c(1,2,3,4,5,6), n, replace = TRUE))
  T[i] <- sum(ifelse(pelemparan_dadu[,1] == pelemparan_dadu[,2],1,0))
}
mean(T)
```

```
## [1] 1.6537
```

dua buah dadu dilempar secara bersamaan sampai muncul sisi angka yang sama. Misalkan T adalah banyaknya pelemparan dadu. Dengan menggunakan simulasi monte carlo estimasi berapa Var[T] ? lakukan simulasi sebanyak 10 ribu kali.

```
set.seed(005)
N <- 10000
n <- 10
T <- rep(NA,N)
for (i in 1:N) {
  pelemparan_dadu <- replicate(2, sample(c(1,2,3,4,5,6), n, replace = TRUE))
  T[i] <- sum((ifelse(pelemparan_dadu[,1] == pelemparan_dadu[,2],1,0)))
}
var(T)
```

```
## [1] 1.379114
```

## halaman 5

Hitung berapa IQR dari nilai mpg ini ?

```
mpg <- c(41.5,50.7,36.6, 37.3, 34.2, 45.0, 48.0, 43.2, 47.7, 42.2,
         43.2, 44.6, 48.4, 46.4, 46.8, 39.2, 37.3, 43.5, 44.3, 43.3)
set.seed(005)
quartil1 <- quantile(mpg, 0.25)
quartil3 <- quantile(mpg, 0.75)
IQR <- quartil3 - quartil1
IQR
```

```
##   75%
## 5.575
```

Tidak ada formula untuk menghitung IQR. Gunakan bootstrap untuk menghitung standard error dari IQR.

```
set.seed(005)
mpg <- c(41.5,50.7,36.6, 37.3, 34.2, 45.0, 48.0, 43.2, 47.7, 42.2,
```

```
        43.2, 44.6, 48.4, 46.4, 46.8, 39.2, 37.3, 43.5, 44.3, 43.3)
n <- length(mpg)
xbar_iqr <- rep(0,n)
for (i in 1:n) {
  x_jack <- mpg[-i]
  xbar_iqr[i] <- sd(x_jack)
}
#standard error jackknife
se_iqr <- sd(xbar_iqr)*(n-1)/sqrt(n)
se_iqr
```

```
## [1] 0.6389609
```

Gunakan metode jacknife untuk menghitung bias dari estimator IQR?

```
set.seed(005)
mpg <- c(41.5,50.7,36.6, 37.3, 34.2, 45.0, 48.0, 43.2, 47.7, 42.2,
        43.2, 44.6, 48.4, 46.4, 46.8, 39.2, 37.3, 43.5, 44.3, 43.3)
n <- length(mpg)
iqr.mpg <- quantile(mpg, 0.75)-quantile(mpg, 0.25)
xbar_iqr <- rep(0,n)
for (i in 1:n) {
  x_iqr <- mpg[-i]
  xbar_iqr[i] <- quantile(mpg, 0.75)-quantile(mpg, 0.25)
}
bias_iqr <- (n - 1)*(mean(xbar_iqr) - mean(iqr.mpg))
bias_iqr
```

```
## [1] 0
```

## halaman 6

Berapa nilai estimasi koefisien determinasi (R2) ?

```
x <- c(4,4,7,7,8,9,10,10,10,11,11,12,12,12,12)
y <- c(2,10,4,22,16,10,18,26,34,17,28,14,20,24,27)
data_frame <- data.frame(x,y)
persamaan.regresi <- lm(y~x, data=data_frame)
rsquare <- summary(persamaan.regresi)$r.squared
rsquare
```

```
## [1] 0.4073877
```

Berikan interpretasi dari R2 di atas?

Keragaman y dapat dijelaskan oleh peubah x yang dimasukan ke dalam model adalah sebesar 40.74% sedangkan sisanya 50.26% dijelaskan oleh peubah lain yang tidak dimasukan kedalam model.

Tuliskan algoritma untuk mengestimasi standard error dari R2 dengan menggunakan bootstrap ?

Estimasi standard error Bootstrap dapat diperoleh dengan algoritma berikut:

- Ambil $B$ sampel Bootstrap independent $X^{*(1)}, \ldots, X^{*(B)}$ dari $F_1$:

$$X_1^{*(b)}, \ldots, X_n^{*(b)}, b = 1, \ldots, B$$

- Evaluasi replikasi Bootstrap:

$$\hat{\theta}^b = t(X^{*(b)}), b = 1, \ldots, B$$

- Estimasi standard error dengan standard deviasi dari $B$ replikasi.

$$\hat{se}_{boot}(\hat{\theta}) = \sqrt{\frac{1}{B-1}\sum_{b=1}^{B}(\hat{\theta}^{*(b)} - \hat{\theta}^{*(\cdot)})^2}$$

dimana $\hat{\theta}^{*(\cdot)} = \frac{1}{B}\sum_{b=1}^{B}\hat{\theta}^{*(b)}$

Berapa estimasi standard error R2 dengan menggunakan metode bootstrap?

```
x <- c(4,4,7,7,8,9,10,10,10,11,11,12,12,12,12)
y <- c(2,10,4,22,16,10,18,26,34,17,28,14,20,24,27)
data_frame <- data.frame(x,y)
persamaan.regresi <- lm(y~x, data=data_frame)
rsquare <- summary(persamaan.regresi)$r.squared
set.seed(005)
n <- length(rsquare)
B <- 1000
xbar.star <- rep(NA, B)
for (i in 1:B) {
  x.star <- sample(x, n, replace = TRUE)
  xbar.star[i] <- mean(x.star)
}
se_r2 <- sd(xbar.star)
se_r2
```

```
## [1] 2.605026
```

Hitung selang kepercayaan bootsrap persentil 90% untuk R2?

```
x <- c(4,4,7,7,8,9,10,10,10,11,11,12,12,12,12)
y <- c(2,10,4,22,16,10,18,26,34,17,28,14,20,24,27)
data_frame <- data.frame(x,y)
persamaan.regresi <- lm(y~x, data=data_frame)
rsquare <- summary(persamaan.regresi)$r.squared
set.seed(005)
n <- length(rsquare)
B <- 1000
delta.star <- rep(NA, B)
for (i in 1:B){
  x.star <- sample(x, n, replace = TRUE)
  xbar.star <- mean(x.star)
  delta.star[i] <- xbar.star - mean(x)
}
d <- quantile(delta.star, c(0.05,0.95))
ci <- mean(x) - c(d[2],d[1])
paste("Selang kepercayaan 90% : [", round(ci[1],4), ",", round(ci[2],4), "]")
```

```
## [1] "Selang kepercayaan 90% : [ 6.5333 , 14.5333 ]"
```

Gunakan standard error bootstrap untuk R2 sebagai estimator statdard error R2. Hitung p-value dari uji hipotesis ini.

```
x <- c(4,4,7,7,8,9,10,10,10,11,11,12,12,12,12)
y <- c(2,10,4,22,16,10,18,26,34,17,28,14,20,24,27)
data_frame <- data.frame(x,y)
persamaan.regresi <- lm(y~x, data=data_frame)
rsquare <- summary(persamaan.regresi)$r.squared
set.seed(005)
B <- 1000
n <- length(rsquare)
T_star <- rep(0, B)
for (i in 1:B) {
  X_star <- sample(rsquare, n , replace = TRUE)
  T_star[i] <- mean(X_star)
}
pval <- sum(T_star <= mean(rsquare)) / B
pval
```

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

Dengan menggunakan taraf signifikan 5%, berikan kesimpulan berdasarkan uji hipotesis di atas.

Berdasarkan uji hipotesis diperoleh nilai p-value = 1 > alfa = 5% maka H0 diterima artinya R2 normal.