

# 1 Exploratory Data Analysis

Vice informací je k dispozici na webové adrese <https://yihui.name/knitr/options/>.

## 1.1 Nactení dat

```
#setwd('/home/pgs/xbendovav/cox')
setwd('C:/Disk D/ND-Skola/02-Vyuka/06-Statistická inference II 2017/17-05-16')
data <- read.delim('15-anova-means-skull.txt', stringsAsFactors=F)
data <- na.omit(data)
data$upface_cat <- car::recode(data$upface.H, " lo:70=1; 70.001:hi=2")
```

```
head(data)
```

```
  id pop sex upface.H upface_cat
1  1  nem  m      73          2
2  2  nem  m      73          2
3  3  nem  m      67          1
4  4  nem  m      75          2
5  5  nem  m      70          1
6  6  nem  m      62          1
```

```
dim(data)
```

```
[1] 163  5
```

## 1.2 Tabulky

	m	s	med	q1	q3
malajska	70.1304	4.9525	70.0000	67.0000	73.0000
peruanska	70.6136	4.1102	70.0000	67.7500	74.0000

Tabulka 1: Tabulka základních charakteristik

Z tabulky základních charakteristik 1 vidíme, že výška horní části tvaru malajské populace je mírně nižší než výška horní části tvaru peruánské populace.

	$\bar{x}$	s	$x_{0.5}$	$x_{0.25}$	$x_{0.75}$
malajska	70.1304	4.9525	70.0000	67.0000	73.0000
peruanska	70.6136	4.1102	70.0000	67.7500	74.0000

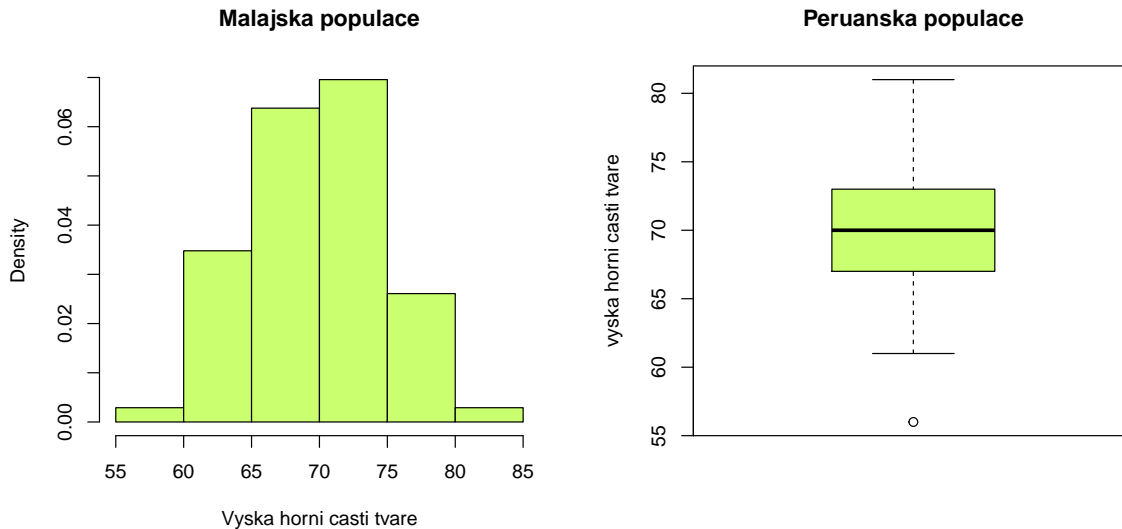
Tabulka 2: Tabulka základních charakteristik

Z tabulky 2 vidíme, že směrodatná odchylka výšky horní části tvaru malajské populace je mírně vyšší než směrodatná odchylka výšky horní části tvaru peruánské populace.

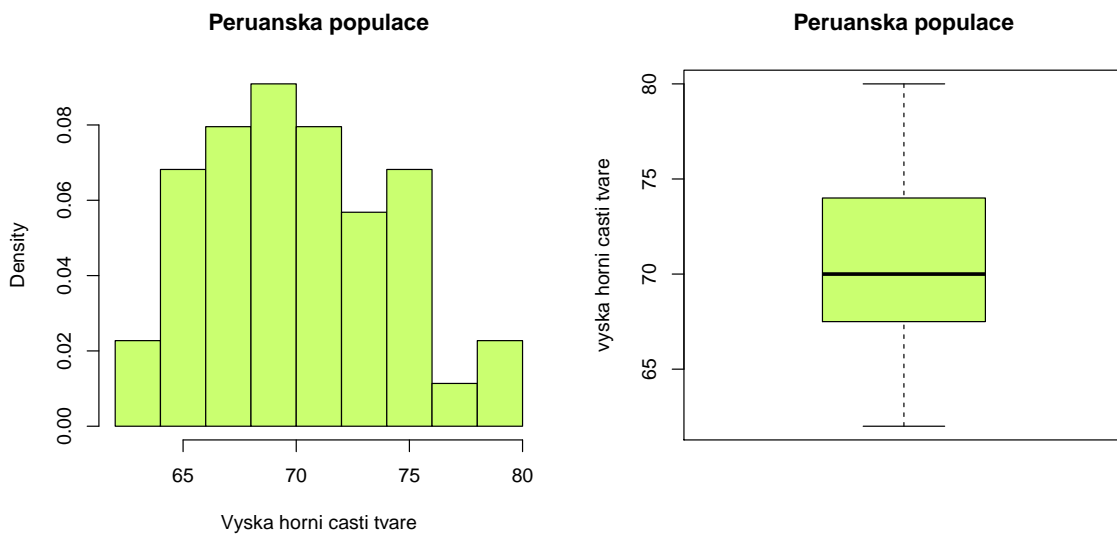
populace	kategorie		$\sum$
	$\leq 70$	$> 70$	
malajska	35	34	69
peruanska	23	21	44
$\sum$	58	55	113

Tabulka 3: Kontingenční tabulka absolutních četností

### 1.3 Grafy, obrázky



Obrázek 1: Vyska horni casti tvare – malajska populace



Obrázek 2: Vyska horni casti tvare – peruanska populace

Z obrázku 1 a 2 je videt, ze rozdeleni horni vysky tvare malajske populace je blize normalnimu rozdeleni nez rozdeleni horni vysky tvare peruanske populace.

## 1.4 Animace

```
like <- function(x,N,p){
  like <- p^x*(1-p)^N
  return(like)
}

monotonnost <- function(N, p1, p2){
  x <- seq(0:N)
  L <- like(x, N, p2)/like(x, N, p1)

  plot(x, L, type='l', ylab='', main='', col='red', lwd=2)
  mtext(expression(L(p[1],x)/L(p[2],x)), side=2, line=2.5)
  mtext('Monotonni pomer verohodnosti', side=3, line=1.5, font = 2)
  mtext(paste('N=', N, ' ; p1=', p1, ' ; p2=', p2, sep=''), line=0.4, cex=0.8)
}

p2 <- 0.5
p1 <- seq(p2-0.01, 0.01, by=-0.01)

for (j in 1:length(p1)) {
  monotonnost(20, p1[j],p2)
}
```

Obrázek 3: Monotonni pomer verohodnosti

Na obrázku 3 je zachycena monotonnost pomeru verohodnosti pro  $X \sim \text{Bin}(N, p)$ .

## 1.5 Mnohonásobne generovani tabulek:

populace	kategorie		$\Sigma$
	$\leq 70$	$> 70$	
malajska	35	34	69
nemecka	6	13	19
$\Sigma$	41	47	88

Tabulka 4: Kontingencni tabulka absolutnich cetnosti – malajska a nemecka populace

populace	kategorie		$\Sigma$
	$\leq 70$	$> 70$	
malajska	8	5	13
cinska	35	34	69
$\Sigma$	43	39	82

Tabulka 5: Kontingencni tabulka absolutnich cetnosti – malajska a cinska populace

populace	kategorie		$\Sigma$
	$\leq 70$	$> 70$	
malajska	7	11	18
bantuska	35	34	69
$\Sigma$	42	45	87

Tabulka 6: Kontingencni tabulka absolutnich cetnosti – malajska a bantuska populace

populace	kategorie		$\Sigma$
	$\leq 70$	$> 70$	
peruanska	6	13	19
nemecka	23	21	44
$\Sigma$	29	34	63

Tabulka 7: Kontingencni tabulka absolutnich cetnosti – peruanska a nemecka populace

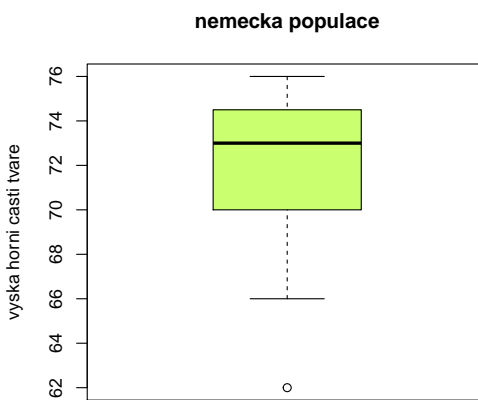
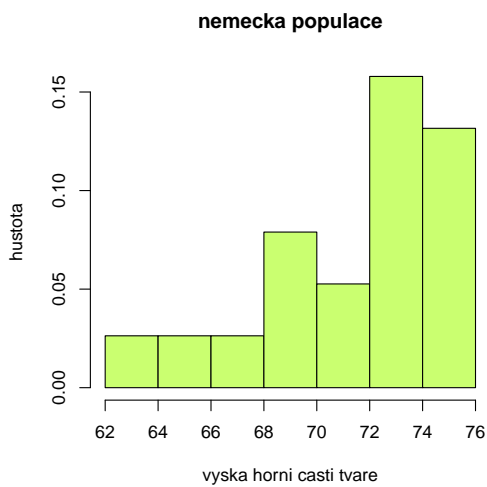
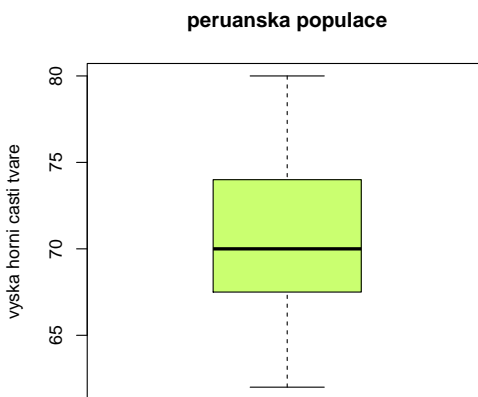
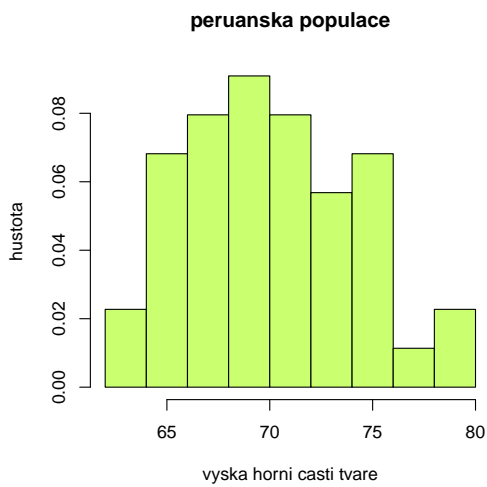
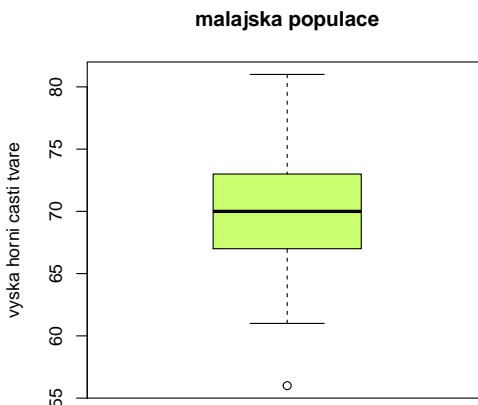
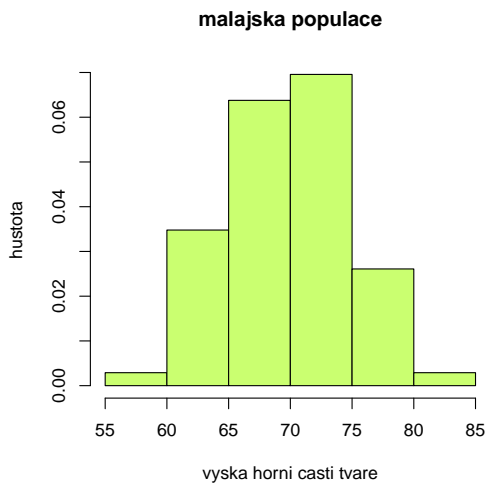
populace	kategorie		$\Sigma$
	$\leq 70$	$> 70$	
peruanska	8	5	13
cinska	23	21	44
$\Sigma$	31	26	57

Tabulka 8: Kontingencni tabulka absolutnich cetnosti – peruanska a cinska populace

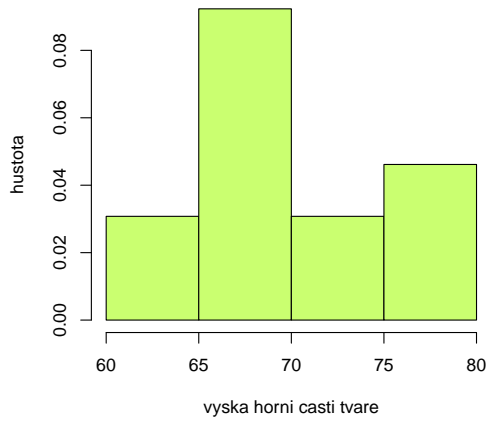
populace	kategorie		$\Sigma$
	$\leq 70$	$> 70$	
peruanska	7	11	18
bantuska	23	21	44
$\Sigma$	30	32	62

Tabulka 9: Kontingencni tabulka absolutnich cetnosti – peruanska a bantuska populace

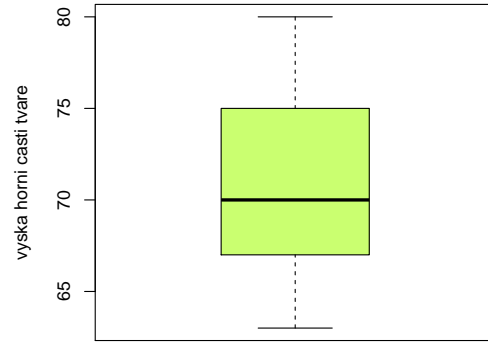
## 1.6 Mnohonásobne generovani grafu:



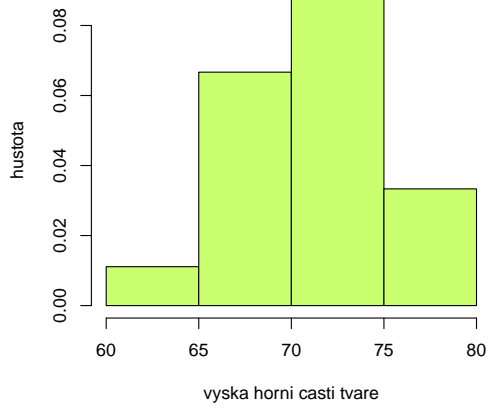
**cinska populace**



**cinska populace**



**bantuska populace**



**bantuska populace**

