

M0130 – 1. PRAKTIKUM : M0130pr01 (*Základy práce s R*)

A. Na adrese

<http://www.math.muni.cz/~kolacek/vyuka/vypsyst/cvr/>

stáhněte všechny cvičební dávky a na základě především prvních třech dávek si procvičte práci se základními datovými strukturami jazyka R. Kromě toho máte možnost čerpat z materiálu, který lze získat na adrese

http://www.math.muni.cz/~kolacek/vyuka/vypsyst/navod_R.pdf.

B. Vybrané matematické a statistické funkce, jednoduché grafické příkazy:

1. Prostředí R umožňuje velmi snadno určovat hodnoty distribučních funkcí, hustot, popř. pravděpodobnostních funkcí pro nejznámější typy rozdělení. Slouží k tomu několik jednoduchých příkazů, začínajících písmeny

- d (*density* – hustoty, popř. pravděpodobnostní funkce),
- p (*probability* – distribuční funkce),
- q (*quantiles* – kvantilová funkce),
- r (*random* – generátory pseudonáhodných čísel)

pro základní typy rozdělení (balíček **stats**) s příkazy

*beta	Beta rozdělení
*binom	binomické rozdělení
*cauchy	Cauchyovo rozdělení
*chisq	(necentraální) χ^2 rozdělení
*exp	exponenciální rozdělení
*f	F-rozdělení
*gamma	Gamma rozdělení
*geom	geometrické rozdělení
*hyper	hypergeometrické rozdělení
*logis	logistické rozdělení
*lnorm	log-normální rozdělení
*multinom	multinomické rozdělení
*nbinom	negativně binomické rozdělení
*norm	normální (Gausovo) rozdělení
*pois	Poissonovo rozdělení
*signrank	rozdělení statistiky <i>Wilcoxon Signed Rank Statistic</i>
*t	Studentovo t-rozdělení
*tukey	rozdělení statistiky <i>Studentized Range</i> (jen p, q)
*unif	rovnorné rozdělení
*weibull	Weibullovo rozdělení
*wilcox	rozdělení statistiky <i>Wilcoxon Rank Sum Statistic</i>

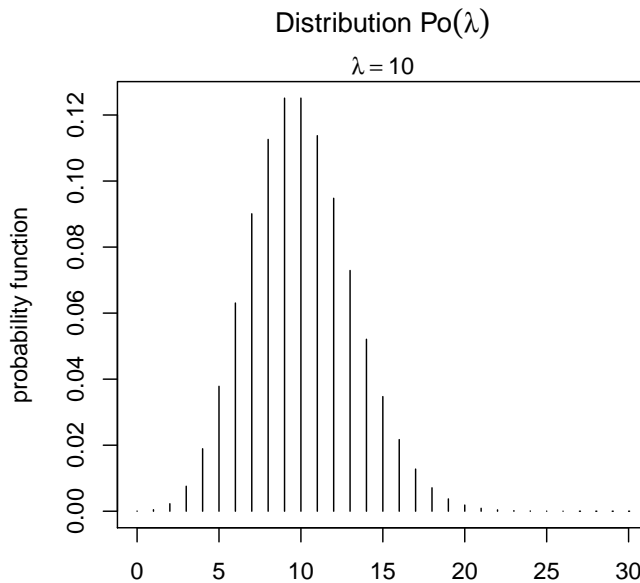
kde * značí počáteční písmeno d, p, q nebo r.

Pokud chceme použít další rozdělení, je třeba nainstalovat speciální balíčky, například `actuar`, `evd`, `mvtnorm`.

PŘÍKLAD 1

Nejprve ukážeme graf pravděpodobnostní funkce diskrétního rozdělení, a to Poissonova. Toto rozdělení má jediný parametr λ , který je zároveň střední hodnotou i rozptylem.

```
> Lambda <- 10
> txtLambda <- substitute(expression(lambda == Lambda), list(Lambda = Lambda))
> plot(0:30, dpois(0:30, Lambda), type = "h", xlab = "", ylab = "probability function",
      main = expression("Distribution Po"(lambda)))
> mtext(eval(txtLambda), line = 0.25)
```



Obrázek 1: Graf pravděpodobnostní funkce Poissonova rozdělení $Po(10)$

PŘÍKLAD 2

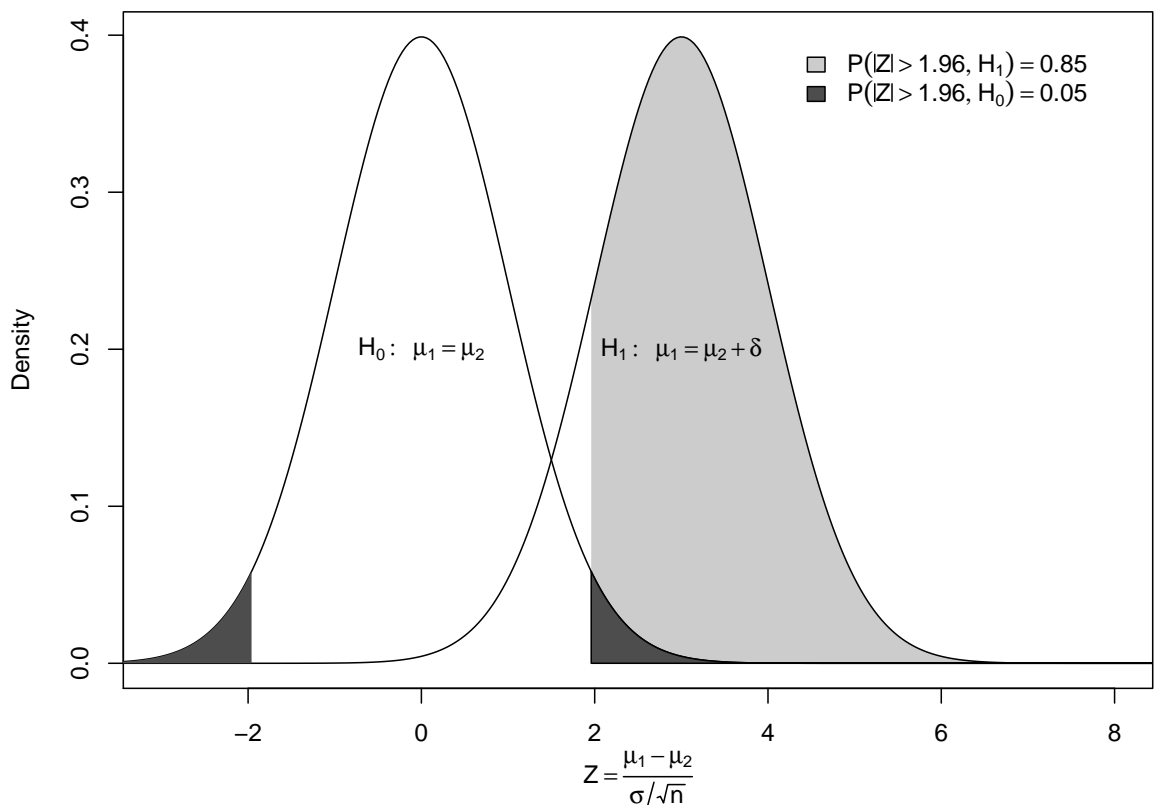
O něco zajímavější je následující příklad, který demonstruje chyby prvního a druhého druhu při testování hypotézy o shodnosti středních hodnot oproti alternativě posunutí (vše za předpokladu normálního rozdělení).

```
> alpha <- 0.05
> x <- seq(-10, 10, length = 500)
> y1 <- dnorm(x)
> y2 <- dnorm(x, m = 3)
> y0 = qnorm(1 - alpha/2)
> p0 <- alpha
> p1 <- 1 - pnorm(y0, m = 3)
> txtLegend <- substitute(expression(P(abs(Z) > y0, H[1]) == p1, P(abs(Z) >
  y0, H[0]) == p0), list(y0 = round(y0, 2), p1 = round(p1, 2), p0 = p0))
> par(mar = c(5, 4, 2, 1))
> plot(x, y2, xlim = c(-3, 8), type = "n", xlab = quote(Z == frac(mu[1] -
  mu[2], sigma/sqrt(n))), ylab = "Density")
> polygon(c(y0, y0, x[x >= y0], x[length(x)]), c(0, dnorm(y0, m = 3),
  y2[x >= y0], 0), col = "grey80", lty = 0)
> lines(x, y2)
> lines(x, y1)
> polygon(c(-y0, -y0, rev(x[x < (-y0)]), x[1]), c(0, dnorm(y0, m = 0),
```

```

      rev(y1[x < (-y0)]), 0), col = "grey30", lty = 0)
> polygon(c(y0, y0, x[x >= y0], 10), c(0, dnorm(y0, m = 0), y1[x >= y0],
      0), col = "grey30")
> legend(4.2, 0.4, fill = c("grey80", "grey30"), legend = eval(txtLegend),
      bty = "n")
> text(0, 0.2, quote(H[0]: ~mu[1] == mu[2]))
> text(3, 0.2, quote(H[1]: ~mu[1] == mu[2] + delta))

```



Obrázek 2: Chyby prvního a druhého druhu

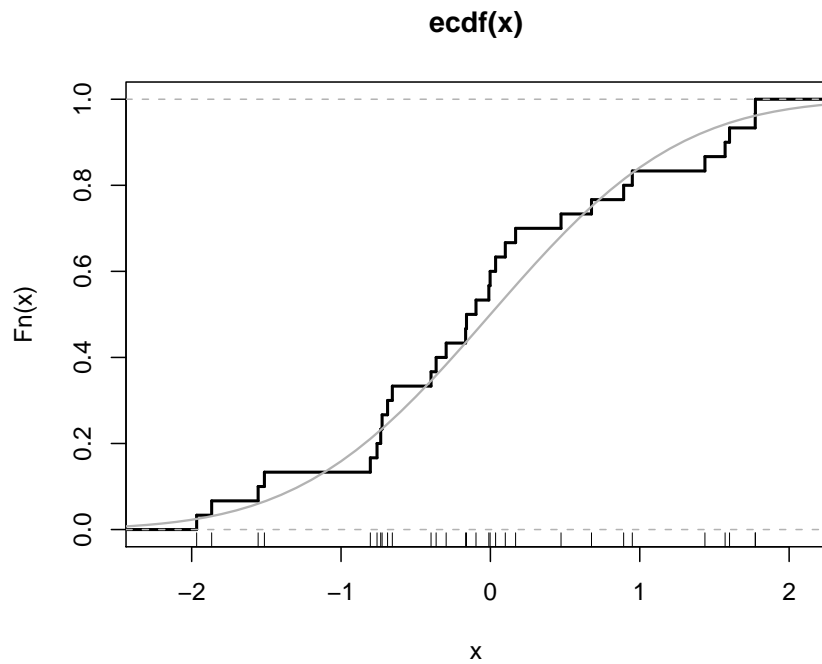
PŘÍKLAD 3

Další příklad zobrazuje empirickou (výběrovou) distribuční funkci spolu s teoretickou distribuční funkcí pro náhodný výběr z normálního rozdělení různých rozsahů.

```

> n = 30
> x <- rnorm(n)
> plot(ecdf(x), verticals = TRUE, do.p = FALSE, lwd = 2)
> curve(pnorm, from = -5, to = 5, add = TRUE, col = "gray70", lwd = 1.5)
> rug(x)

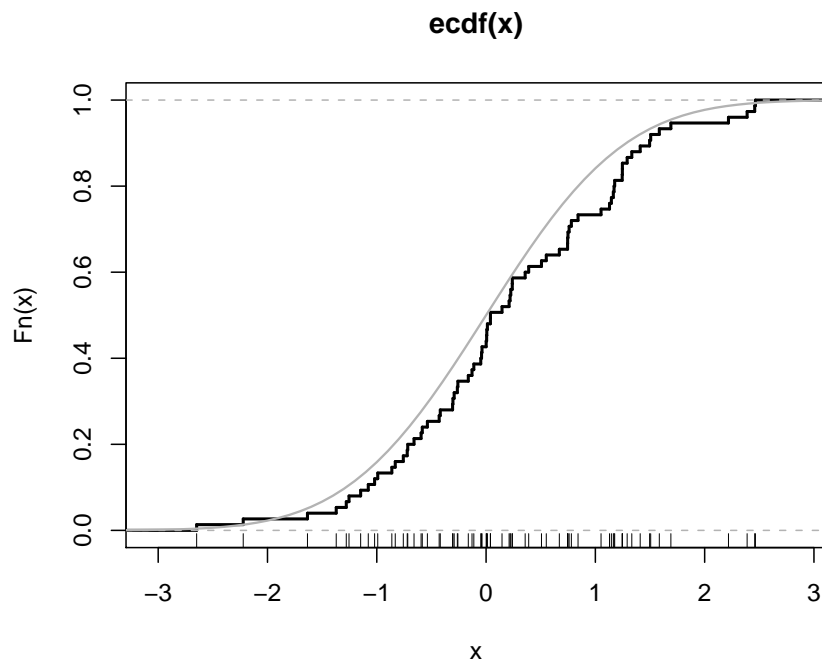
```

Obrázek 3: Empirická distribuční funkce pro normální náhodný výběr rozsahu $n = 30$

```

> n = 75
> x <- rnorm(n)
> plot(ecdf(x), verticals = TRUE, do.p = FALSE, lwd = 2)
> curve(pnorm, from = -5, to = 5, add = TRUE, col = "gray70", lwd = 1.5)
> rug(x)

```

Obrázek 4: Empirická distribuční funkce pro normální náhodný výběr rozsahu $n = 75$

ÚKOL 1

Graficky porovnej hustotu (distribuční funkci) Studentova rozdělení s 3 stupni volnosti se standardizovaným normálním rozdělením.

ÚKOL 2

Porovnej hustoty exponenciálního rozdělení se středními hodnotami 1 a 5.