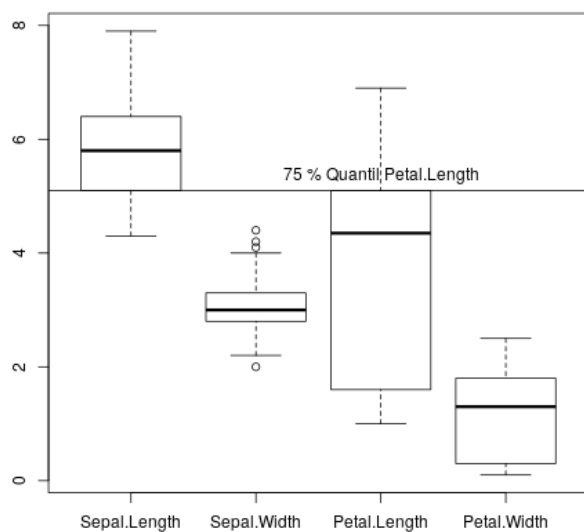


10 Mehr zu Graphik in R

Aufgabe 1:

Erzeugen Sie den in der Datei `boxplot.png` abgebildeten Plot und speichern Sie ihn ebenfalls in eine Datei `boxplot.png`. Dafür wurde der Datensatz `iris` verwendet.



Aufgabe 2:

Laden Sie den Datensatz `trees` aus dem Paket `datasets`

- Visualisieren Sie die Daten in paarweisen Streudiagrammen.
- Erzeugen Sie nun 2 Streudiagramme nebeneinander, wo Sie jeweils für die Hälfte der kleinsten sowie der größten Bäume Volumen (`Volume`) gegen Umfang (`Girth`) auftragen. Speichern Sie die Graphik in einer PDF Datei. Achten Sie darauf, dass die x- und y-Achsen gleich sind und beschriften Sie die Graphik geeignet.

Aufgabe 3:

Laden Sie den Datensatz `UCBAdmissions`. Der Datensatz soll analysiert werden, um festzustellen, ob Frauen bei der Aufnahme an der UC Berkeley benachteiligt wurden. Der Verdacht, dass Frauen benachteiligt wurden, ist aufgekommen, da der Anteil der Frauen, die aufgenommen wurden, bzgl. aller Frauen, die sich beworben haben, niedriger war als der entsprechende Anteil bei den Männern.

- Visualisieren Sie den Datensatz geeignet in einem Mosaicplot, sodass man den Anteil der aufgenommenen Männern mit dem der aufgenommenen Frauen innerhalb eines Departments gut vergleichen kann.

- Berechnen Sie die relativen Aufnahmehäufigkeiten der Männer und Frauen jeweils für jedes Department separat und die insgesamt relativen Aufnahmehäufigkeiten bei Männern und Frauen (über alle Departments).
- Wie sind die relativen Aufnahmehäufigkeiten für die einzelnen Departments und insgesamt? Wie hoch ist der Frauenanteil bei den Bewerbern bei den einzelnen Departments und insgesamt?

Aufgabe 4:

Schreiben Sie eine Funktion, die die Dichte und Verteilungsfunktion einer gegebenen Verteilung für gegebene Parameter jeweils in einer Graphik darstellt.

```
> visualize_dist <- function(name, dist_par, x, discrete = FALSE, ...) {}
```

Die Argumente sollen folgendermaßen definiert sein:

- **name**: Name der Verteilung, z.B., "norm".
- **dist_par**: Benannte Liste mit den Parametern der Verteilung, die spezifiziert werden.
- **x**: Werte, an denen die Dichte und die Verteilungsfunktion ausgewertet werden sollen.
- **discrete**: Logischer Wert, ob die Verteilung diskret oder stetig ist. Im diskreten Fall soll die Dichte ein Stabdiagramm und die Verteilungsfunktion eine Treppenfunktion, während im stetigen Fall diese für geeignet dichte x-Werte einfach linear interpoliert werden können.

Hinweis: Der Type der Graphik kann mit dem Argument **type** beim Aufruf von **plot()** verändert werden. D.h. eine Stufenfunktion bekommt man mithilfe von **type = "s"**, eine linear interpolierte mit **type = "l"** und eine Stabdiagramm mit **type = "h"**.

Mithilfe von **par(ask = TRUE)** kann man sicherstellen, dass man im interaktivem Modus von R gefragt wird, bevor ein Graphikfenster geschlossen und ein neues geöffnet wird.

Aufgabe 5:

Mithilfe der Inversionsmethode können Zufallszahlen aus einer Verteilung F gezogen werden, indem man erst auf $[0, 1]$ gleichverteilte Zufallszahlen erzeugt und diese dann mithilfe der inversen Verteilungsfunktion (bzw. Quantilsfunktion) transformiert. D.h.

$$u \sim U[0, 1]$$

$$x = F^{-1}(u)$$

ergibt, dass $x \sim F$.

Erzeugen Sie mithilfe der Inversionsmethode Zufallszahlen aus einer trunkeierten Normalverteilung. Schreiben Sie dafür die folgende Funktion

```
> rtrnorm <- function(n, mean = 0, sigma = 1, lower = -Inf, upper = Inf) {}
```

Dabei gilt:

- **mean**: Erwartungswert der Normalverteilung vor der Trunkierung
- **sigma**: Standardabweichung der Normalverteilung vor der Trunkierung
- **lower, upper**: obere und untere Schranken des Wertebereichs

Wenden Sie die Funktion folgendermaßen an:

```
> x <- rtrnorm(1000, 3, 1, -5, 4)
```

Visualisieren Sie die Stichprobe und vergleichen Sie deren Verteilung mit der theoretischen.