Übungen zur Vorlesung „ Statistik “

Abgabetermin: Montag, 19. November 01, in der Vorlesung

Aufgabe 13 : **a)** Es sei $\mathfrak{Z}$ standard-normalverteilt auf $\mathbb{R}^n$, und $b \in \mathbb{R}^n$. Berechnen Sie mit Fishers Lemma die Verteilung des Skalarprodukts von $\mathfrak{Z}$ und b.
b) X_i sei $\mathfrak{N}(\mu_i, \sigma_i^2)$-verteilt ($i = 1, 2$), und X_1, X_2 seien unabhängig. Ermitteln Sie unter Verwendung von a) die Verteilung von $X_1 + X_2$.

Aufgabe 14 : *"Wie erzeugt man normalverteilte Zufallsvariable aus gleichverteilten?"*
U und V seien unabhängig und gleichverteilt auf dem Einheitsintervall. Zeigen Sie Sie, dass $\sqrt{-2\log U}(\cos(2\pi V), \sin(2\pi V))$ standard-normalverteilt auf $\mathbb{R}^2$ ist.

Aufgabe 15 : (Aus: J.A.Rice „Mathematical Statistics and Data Analysis“, 2. Aufl., Duxbury Press, 1995)
A study was done to compare the performances of engine bearings made of different compounds (J. McCool, Analysis of single classification experiments based on censored samples from the two-parameter Weibull distribution, J. Statist. Planning and Inference 3 (1979), 39-68). Ten bearings of each type were tested. The following table gives the times until failures (in units of millions of cycles):

Type I	Type II
3.03	3.19
5.53	4.26
5.60	4.47
9.30	4.53
9.92	4.67
12.51	4.69
12.95	12.78
15.21	6.79
16.04	9.37
16.84	12.75

a) Use normal theory to test the hypothesis that there is no difference between the two types of bearings.
b) Test the same hypothesis using a nonparametric method.
c) Which of the two methods – that of part (a) or of part (b) – do you think ist better in this case?
d) Estimate π, the probability that a type I bearing will outlast a type II bearing.
e) Use the bootstrap to estimate the sampling distribution of $\hat{\pi}$ and its standard error.
f) Use the bootstrap to find an approximate 90% confidence interval for π.

Aufgabe 16 : $\mathfrak{X}$ sei eine zufällige Stichprobe des Umfangs n aus einer Verteilung ν auf $\mathbb{R}$; μ sei der Erwartungswert und σ^2 die Varianz von ν. Wir wissen, dass $T := \sqrt{n}\ (\text{mean}(\mathfrak{X}) - \mu)/\text{sd}\ (\mathfrak{X})$ Student$(n-1)$ verteilt ist, falls ν eine Normalverteilung ist. In der Praxis wird eine "approximative Student$(n-1)$-Verteiltheit" von T oft auch dann unterstellt, wenn man nicht davon ausgehen kann, dass ν eine Normalverteilung ist. Eine Begründung, die man manchmal hört, ist die folgende: "Zumindest für große n ist (wegen des Gesetzes der großen Zahlen) sd $(\mathfrak{X})$ annähernd gleich σ, und (wegen des Zentralen Grenzwertsatzes) die Verteilung von $\sqrt{n}\ (\text{mean}(\mathfrak{X})-\mu)/\sigma$ annähernd standard-normal. Also ist die Verteilung von T für große n annähernd standard-normal. Weil auch Student$(n-1)$ für große n annähernd standard-normal ist, ist die Verteilung von T für große n annähernd gleich Student$(n-1)$."
Wahr, schön und gut - im Grenzwert $n \to \infty$ können wir also von vornherein die Verteilung von T durch die Standard-Normalverteilung ersetzen. Wie sieht es aber für kleinere n aus?
Machen Sie sich in ein paar Fällen ein Bild, indem Sie Quantilplots der (empirischen) Verteilung von T (mit 1000 oder, wenns Ihr Computer aushält, mit 10000 Replikationen) einmal gegen Student$(n-1)$ und einmal gegen $\mathfrak{N}(0,1)$ herstellen und Ihre Lieblingsquantile vergleichen. Geben Sie einen Befund!

Vorschlag: a) ν = Uniform([0, 1]) b) ν = Exp(1) c) $\nu = (1/4)\mathfrak{N}(-2,1) + (3/4)\mathfrak{N}(2,1)$, und (als Check) d) $\nu = \mathfrak{N}(0,1)$, jeweils mit $n = 10$ und $n = 20$.

So hab ich's mit R gemacht:

```
T← function(x,mu, n) sqrt(n)*(mean(x)-mu)/sd(x)
n←10
mu← 0.5
v←c()
for(i in 1:1000) {x← runif(n) ; v[i]←T(x,mu ,n)}
plot(qt(ppoints(v), n-1), sort(v), abline(0,1))
qqnorm(v)
abline(0,1)
qnorm(0.025); qnorm(0.975)
qt(0.025, n-1); qt(0.975, n-1)
quantile(v, 0.025); quantile(v, 0.975)
```