Statistik 5

Übung 6 kleine Ergänzung

Einleitung und Ziele der Übung

Dieser Teil der Übung dient als Ergänzung zu Übung 6. Dort hatten Sie eine Funktion zum Bootstrappen des t-Tests bei abhängigen Stichproben erstellt. Wir wollen jetzt noch 1 Funktion erstellen um dafür eine a priori Poweranalyse durchführen zu können.

1. A priori - Einführung

Für eine a priori Poweranalyse benötigen wir immer die Effektstärke, das Signifikanzniveau und die gewünschte Teststärke. Die Effektstärke berechnet sich aus $\frac{\bar{x}_d}{s_d}$, wobei $\bar{x}_d$ der Mittelwert der Differenzen und s_d die Standardabweichung der Differenzen ist. Die Übung wird Sie jetzt Schritt für Schritt bis zur Fertigstellung der Funktion führen.

1. A priori - Übung 1

Zunächst erstellen wir die Basis unserer Funktion Apriori, sie soll als Argumente die Effektstaerke, das Signifikanzniveau, die Teststärke und die Richtung unserer Hypothese erhalten. Legen Sie innerhalb der Funktion eine Variable nStichprobe an, welcher Sie den Wert 0 zuweisen und geben Sie nStichprobe anschliessend zurück.

1. A priori - Übung 2

Ihre Funktion sollte jetzt ungefähr so aussehen:

```
Apriori <- function (Effektstaerke, Signifikanzniveau, Teststaerke, Richtung){
  nStichprobe <- 0
  # Hier arbeiten wir weiter
  return(nStichprobe)
}
```

Jetzt überlegen wir uns zunächst wie wir die Stichproben simulieren können (T1 und T2). Wir wissen:

- $Effektstaerke = \frac{\bar{x}_d}{s_d}$, da uns für die Simulation die Werte eigentlich egal sind (solange die Normalverteilungsannahme für die Differenzen gegeben ist), gehen alle Stichproben für die $Effektstaerke * s_d = \bar{x}_d$ gilt.
- das bedeutet T1 können beliebige Werte sein, dann simulieren wir die Differenzen zu T1, addieren diese zu T1 und bestimmen dadurch die Werte von T2.
- Wir legen (willkürlich) fest $T1 ~ (0,1)
- dann gilt $T2 = T1 + (Effektstaerke,1) (der Mittelwert der Unterschiede ist Effekstaerke * zugehörige Standardabweichung)

Generieren Sie jetzt T1 und T2 in Ihrer Funktion mit einer Stichprobengrösse von 10 und führen Sie damit tStat_AS_Boot mit 500 Bootstraps aus. Sperichern Sie das Ergebnis in der Variable signifikant.

1. A priori - Übung 3

Ihr Code sollte jetzt ungefähr so aussehen:

```
Apriori <- function (Effektstaerke, Signifikanzniveau, Teststaerke, Richtung){
  nStichprobe <- 0

  T1 <- rnorm(10,0,1)
  T2 <- T1 + rnorm(10,Effektstaerke,1)
  Signifikant <- tStat_AS_Boot(T1, T2, 500, Richtung, Signifikanzniveau)

  return(nStichprobe)
}
```

Jetzt haben wir das ganze für eine Stichprobe gemacht. Allerdings brauchen wir bei der Poweranalyse durch Simulation viele Stichproben. Es wäre also praktisch, wenn wir Stichprobenerstellung und Signifikanzprüfung in eine eigene Funktion Signifikanz() mit den Argumenten Effektstaerke, Signifikanzniveau, Stichprobenumfang, Bootstraps und Richtung auslagern. Diese Funktion gibt uns dann das Ergebnis von t_Stat_AS_Boot zurück. Damit können wir die FUnktion innerhalb von Apriori mit replicate nutzen, um die nötige Anzahl an Stichproben zu generieren.

Erstellen Sie die Funktion Signifikanz(). Achten Sie darauf innerhalb der Funktion Signifikanz() den Stichprobenumfang nun durch das zugehörige Argument festzulegen, genauso wie die Anzahl an Bootstraps. und ersetzen Sie Ihren Code in der Funktion Apriori durch ein replicate im Umfang von 1000 der Funktion Signifikanz().

1. A priori - Übung 4

Ihr Code sollte jetzt ungefähr so aussehen:

```
Signifikanz <- function(Effektstaerke, Signifikanzniveau, Stichprobenumfang, Bootstraps, R
  T1 <- rnorm(Stichprobenumfang,0,1)
  T2 <- T1 + rnorm(Stichprobenumfang,Effektstaerke,1)
  Signifikant <- tStat_AS_Boot(T1, T2, Bootstraps, Richtung, Signifikanzniveau)
  return(Signifikant)
}

Apriori <- function (Effektstaerke, Signifikanzniveau, Teststaerke, Richtung){
  nStichprobe <- 0
  Signifikant <- replicate(1000, Signifikanz(Effektstaerke,Signifikanzniveau,10,500, Richt
  return(nStichprobe)
}
```

Wir haben ein Problem. Durch Simulationen können wir immer für ein bestimmtes n die Testärke bestimmen. Das passiert wenn wir sum(Signifikant) durch (in dem Fall) 1000 rechnen.

Wir haben die Teststärke in Abhängigkeit von n. Wir brauchen aber n in Abhängigkeit der Teststärke. Um das zu lösen müssen wir also so lange Stichprobengrössen ausprobieren, bis wir eine Stichprobengrösse finden, deren Teststärke gleich oder minimal grösser als unsere vorgegebene Teststärke ist.

-Fortsetzung nächste Seite-

1. A priori - Übung 5

Wir brauchen eine Schleife! In dieser gehen wir systematisch Stichprobengrössen durch (von klein bis gross), wir eine Stichprobengrösse mit ausreichend Teststärke gefunden haben. Wenn wir allerdings bei sehr kleinen Zahlen anfangen (z.B. $n = 5$) und die richtige Zahl aber erst bei 400 ist, braucht es natürlich sehr lange. Daher wollen wir zunächst zwei neue Argumente zu unserer Funktion Apriori hinzufügen:

- nStart (kleinste Stichprobe, die getestet werden soll)
- nSchritt (um welchen Wert wird die Stichprobengrösse in jedem Schleifendurchlauf erhöht)

Legen Sie zusätzlich zwei neue Variablen aktTeststaerke und aktStichprobe an (direkt unter nStichprobe) und weisen sie der ersten den Wert 0 und der zweiten den Wert nStart zu.

Weisen Sie aktTestaerke vor der return-Zeile den Wert sum(Signifikant)/1000 zu.

1. A priori - Übung 6

Ihr Code sollte jetzt ungefähr so aussehen:

```
#An der Funktion Signifikanz ändert sich nichts mehr, daher weggelassen
Apriori <- function (Effektstaerke, Signifikanzniveau, Teststaerke, Richtung, nStart, nSch
  nStichprobe <- 0
  aktTeststaerke <- 0
  aktStichprobe <- nStart
  Signifikant <- replicate(1000, Signifikanz(Effektstaerke,Signifikanzniveau, 10, 500, Ric
  aktTeststaerke <- sum(Signifikant)/1000
  return(nStichprobe)
}
```

Jetzt brauchen wir eine Schleife, die…

- eine while Schleife ist (wir wissen nicht wie viele Durchgaenge wir brauchen)
- wiederholt, solange die aktTesstaerke kleiner als die geforderte Teststaerke ist
- für die aktuelle Stichprobengroesse die Zeilen Signifikant <- … und aktTestaerke <-… ausführt (also nicht immer 10!)
- anschliessend die aktuelle Stichprobengroesse um nSchritt erhöht (danach beginnt der nächste Durchgang)

Nach der Schleife weisen wir nStichprobe den Wert aktStichprobe - nSchritt zu (wir haben einmal zu viel hochgezählt). Setzen Sie dies um und Ihre Funktion ist fertig!

1. A priori - Übung 7

```r
Apriori <- function (Effektstaerke, Signifikanzniveau, Teststaerke, Richtung, nStart, nSch
  nStichprobe <- 0
  aktTeststaerke <- 0
  aktStichprobe <- nStart
  while (aktTeststaerke < Teststaerke){
    Signifikant <- replicate(1000, Signifikanz(Effektstaerke,Signifikanzniveau, aktStichpr
    print(aktStichprobe) #Ergänzt, damit Sie den Arbeitsfortschritt sehen
    aktTeststaerke <- sum(Signifikant)/1000
    print(aktTeststaerke) #Ergänzt, damit Sie den Arbeitsfortschritt sehen
    aktStichprobe <- aktStichprobe + nSchritt
  }
  nStichprobe <- aktStichprobe - nSchritt
  return(nStichprobe)
}
```

Probieren Sie die folgenden Varianten aus:

```r
Apriori(0.5, .05, .70, "u", 15, 2)
Apriori(0.1, .05, .70, "u", 100, 100)
# Natürlich können Sie nach einer groben Schätzung eine genauere folgenlassen und dabei ev
# Beachen Sie einfach, für jedes n werden jetzt gerade schon 1000*500 Werte berechnet
#Bei der letzten Durchführung könnten Sie zum Beispiel in einem nächsten Schritt,
#ab 580 starten und in 10er Schritten vorgehen
Apriori(0.1, .05, .70, "u", 580, 10)
```