Statistik 5

Übung 4

Einleitung und Ziele der Übung

Herzlich Willkommen zur Übung Statistik 5. Bei dieser Übung wollen wir *t*-Tests für unabhängige Stichproben mit passenden Konfidenzintervallen und *U*-Tests vergleichen.

Agenda

- Konfidenzintervall Mittelwertsunterschiede
- *t*-Test für unabhängige Stichproben
- *U*-Test bzw. Wilcoxon-Rangsummentest

1. Konfidenzintervall Mittelwertsunterschiede - Einführung

In der Vorlesung haben wir uns bereits die Idee des Signifikanztests angeschaut. In dieser Übung soll geschaut werden, wie sich *t*-Test, Konfidenzintervalle und *U*-Test verhalten.

1. Konfidenzintervall Mittelwertsunterschiede - Übung 1

Dafür sollen jeweils 100’000 Stichproben aus zwei Populationen (A, B) gezogen werden. Mit diesen werden wir durchweg arbeiten:

- Jede Population besteht aus je zwei Subpopulationen 1 und 2, wobei für jede Teilpopulation in jeder Stichprobe $n = 64$ gelten soll.
- Die Populationsparameter für A1 seien: $\mu_{A1} = 100, \quad \sigma^2_{A1} = 144$
- Die Populationsparameter für A2 seien: $\mu_{A2} = 100, \quad \sigma^2_{A2} = 144$

} nicht sign

α = Fehler 1 Art.

- Die Populationsparameter für B1 seien: $\mu_{B1} = 100, \quad \sigma^2_{B1} = 144$
- Die Populationsparameter für B2 seien: $\mu_{B2} = 106, \quad \sigma^2_{B2} = 144$

} sign.

β = Fehler 2 Art.

Benutzen Sie den Seed 1000. Simulieren Sie die die 100’000 Stichproben je Subpopulation und speichern Sie das Ergebnis jeweils in einer Variable nach dem Schema: Population_Subpopulation, z.B. A_1

Wenn Sie alles richtig gemacht haben, sollten Sie etwa folgende Werte für die jeweils erste Person ihrer Stichprobe haben: 94.7, 110.0, 99.6, 89.4

$\Delta = A_1 - A_2$ -> kein Unterschied
$\Delta = B_1 - B_2$ -> ein Unterschied

α = ist nein, wird aber als ja genommen
β = ist ja, wird aber als nein genommen

1. Konfidenzintervall Mittelwertsunterschiede - Übung 2

Berechnen sie für jede der Stichprobe und jede Subpopulation die folgenden Stichprobenkennwerte und speichern Sie diese in jeweils in Variable. Sie können dafür wieder apply verwenden.

- $\overline{x}$ (z.B. A_1_M als Variable)
- s (z.B. A_1_SD als Variable)

Haben Sie alles richtig gemacht, sollten Sie jetzt 8 Variablen mit jeweils 100’000 Werten haben.

Berechnen Sie jetzt folgende Grössen für den Vergleich der Subgruppen 1 und 2 wieder getrennt für die Populationen A und B. Sie sollten vier neue Variablen (A_sigma2_inn, B_sigma2_inn, MDiff_Sigma2_A, MDiff_Sigma2_B) mit jeweils 100’000 Werten haben. Sie brauchen diesmal kein apply () oder Schleifen.

$$\widehat{\sigma}^2_{inn} = \frac{(n_1 - 1) * s_1^2 + (n_2 - 1) * s_2^2}{n_1 + n_2 - 2}$$

$$\widehat{\sigma}_{\overline{x}_1 - \overline{x}_2} = \sqrt{\frac{\widehat{\sigma}^2_{inn}}{n_1} + \frac{\widehat{\sigma}^2_{inn}}{n_2}}$$

1. Konfidenzintervall Mittelwertsunterschiede - Übung 3

Bestimmen Sie den benötigten *t*-Wert für ein 95%-Konfidenzintervall für die Mittelwertsunterschied und speichern Sie ihn in der Variable t.

Prüfen Sie Ihren Wert hier (2 Nachkommastellen):

Bestimmen Sie jetzt die Ober- und Untergrenzen ihrer Konfidenzintervalle für jede Stichprobe und speichern Sie die Ergebnisse in den Variablen A_ug, A_og, B_ug, B_og.

Prüfen Sie für jedes Konfidenzintervall ob die 0 ausserhalb der Ober- und Untergrenzen liegt und speichern Sie das Ergebnis in den Variablen A_Ausserhalb, B_Ausserhalb. Sie benötigen keine if-Bedingung aber Vergleiche und den oder-Operator |.

Beantworten Sie folgende Fragen:

- Wie oft lag für Stichprobe A die 0 ausserhalb der Konfidenzintervalle?
- Wie oft lag für Stichprobe B die 0 ausserhalb der Konfidenzintervalle? X/100'000
- Wie oft lag für Stichprobe A die 0 innerhalb der Konfidenzintervalle?
- Wie oft lag für Stichprobe B die 0 innerhalb der Konfidenzintervalle? B/100'000

Poor

OK-Fehler weil es gibt kein Unterschied

B-Fehler weil es gibt einen

1. Konfidenzintervall Mittelwertsunterschiede - Übung 4

a) Was bedeutet es, wenn die 0 innerhalb oder ausserhalb des Konfidenzintervalls liegt?

B) Wir hatten eine Konfidenz von 95% gewollt. Passen unsere Ergebnisse zu diesem Wert? Begründen Sie Ihre Antwort.

C) Population B wurde so gewählt, dass unsere Stichprobe mit einer gewissen Power/Teststärke ein signifikantes Ergebnis liefern. Wie gross sind Fehler 2. Art und die Teststärke für die Stichproben von Population B?

B) wenn es 95% ist sollten 5% der Daten ausserhalb liegen

2. *t*-Test für unabhängige Stichproben - Einführung

Die Statistik für den *t*-Test für unabhängige Stichproben ist definiert als:

$$t = \frac{\overline{x}_1 - \overline{x}_2}{\widehat{\sigma}_{\overline{x}_1 - \overline{x}_2}}$$

Sie gibt also an, wie gross das Verhältnis zwischen der Mittelwertsdifferenz zur geschätzten Streuung der Mittelwerte (über viele theoretische Stichproben) ist. Je grösser das Verhältnis, desto mehr können wir annehmen, dass der beobachtete Unterschied nicht nur Zufall war.

Erkenn Sie auch, dass diese Formel der umgestellten Formel für das Konfidenzintervall entspricht, ausser das hier die empirische *t*-Wert berechnet wird, während in die Gleichung des Konfidenzintervalls der kritische *t*-Wert eingesetzt wird.

2. *t*-Test für unabhängige Stichproben - Übung 1

Sie haben eigentlich schon alles, um die empirischen *t*-Werte zu bestimmen. Bitte bestimmen Sie jetzt die empirischen *t*-Werte für beide Populationen und speichern Sie die Ergebnisse in A_t und B_t.

Für die Prüfung der statistischen Signifikanz könnten Sie jetzt einfach die empirischen *t*-Werte mit dem kritischen *t*-Wert von zuvor vergleichen.

Wir wollen einen anderen Weg gehen. Bestimmen Sie für jeden *t*-Wert den zugehörigen *p*-Wert mit der Funktion pt(). Probieren Sie die Funktion am besten mit dem kritischen *t*-Wert von zuvor aus. Sie werden das erste, das zweite und das vierte Argument der Funktion brauchen. Denken Sie auch daran, dass der *p*-Wert für eine ungerichtete Hypothese wäre. Ein letzter Hinweis, Sie müssen (um die Richtung ignorieren zu können) die Beträge der *t*-Werte mittels abs() nehmen.

Sobald Sie die Funktion verstanden haben, bestimmen sie die *p*-Werte und speichern Sie die Ergebnisse in A_p und B_p.

Beantworten Sie folgende Fragen:

2. *t*-Test für unabhängige Stichproben - Übung 2

Wir wissen, dass eigentlich Subpopulation B2 einen höheren Mittelwert in der Population aufweist, als Subpopulation B1. Wie oft haben Sie ein signifikantes Ergebnis erzielt, wobei in der Stichprobe die falsche Gruppe den höheren Mittelwert hatte?

Sie können diese Frage mit den Konfidenzintervallen oder den *t*-Tests lösen.

Wir haben jetzt gesehen, dass Konfidenzintervalle und die zugehörigen parametrischen Tests exakt die gleichen Ergebnisse erzielen.

Jetzt wollen wir uns noch die *p*-Werte für die nicht signifikanten Ergebnisse anschauen. Bitte erstellen sie 6 Histogramme für die *p*-Werte von

- Population A
- Population B
- Population A, aber nur die nicht signifikanten *p*-Werte
- Population B, aber nur die nicht signifikanten *p*-Werte
- Population A, aber nur die signifikanten *p*-Werte
- Population B, aber nur die signifikanten *p*-Werte

2. *t*-Test für unabhängige Stichproben - Übung 3

Was ist Ihnen bei den vorherigen Histogrammen aufgefallen?

Schauen wir uns noch die Mittelwertsunterschiede in Bezug auf die zugehörigen *p*-Werte an. Erstellen Sie dazu passende Streudiagramme für

- Population A
- Population B
- Population A, aber nur Stichproben mit nicht signifikanten *p*-Werten
- Population B, aber nur nur Stichproben mit nicht signifikanten *p*-Werten
- Population A, aber nur Stichproben mit signifikanten *p*-Werten
- Population B, aber nur Stichproben mit signifikanten *p*-Werten

Was fällt Ihnen auf?

3. *U*-Test - Einführung

Jetzt wollen wir noch den *U*-Test mit dem *t*-Test vergleichen. Der *U*-Test wurde wie folgt durchgeführt:

1. Ränge zuweisen über beide Gruppen hinweg
2. Rangsummen für beide Gruppen berechnen
3. Für jede Gruppe einen *U*-Wert berechnen
4. Das Minimum der beiden U-Werte ist die Teststatistik (der empirische *U*-Wert)
5. Da R die Wilcoxon-Verteilung statt der *U*-Verteilung verwendet, nehmen wir stattdessen die Rangsumme von Gruppe 1 und ziehen die minimal mögliche Rangsumme für diese Gruppe ab.Die beiden Tests sind aber äquivalent.

$$U_1 = n_1 * n_2 + \frac{n_1 * (n_1 + 1)}{2} - R_1$$

$$U_2 = n_1 * n_2 + \frac{n_2 * (n_2 + 1)}{2} - R_2$$

$$W = R_1 - \frac{n_1 * (n_1 + 1)}{2}$$

3. *U*-Test - Übung 1

Da die Rangbestimmung etwas kompliziert ist nutzen Sie bitte zunächst folgenden Code, um die Ränge für die jeweilige Subpopulation 1 zu ermitteln:

```
n <- 64
A_1_Rang <- apply(rbind(A_1, A_2),2,rank)[1:n,]
B_1_Rang <- apply(rbind(B_1, B_2),2,rank)[1:n,]
```

Nachdem Sie die Ränge wie oben ermittelt haben, bestimmen Sie die Rangsumme je Stichprobe für Gruppe 1 und ziehen Sie davon die theoretisch niedrigste mögliche Rangsumme ab. Speichern Sie Ihre Ergebnisse in A_W und B_W.

3. *U*-Test - Übung 2

Der Umgang mit pwilcox ist etwas kompliziert, da die Verteilung nicht um 0 symmetrisch ist sondern der Erwartungswert unter der Nullhypothese bei $(n_1 * n_2)/2$ liegt. Daher hier der Code um die *p*-Werte zu bestimmen

```
# Da n1 = n2 ist gilt:
Erwartungswert <- 64^2/2
# normalerweise bekommt pwilcox, als Argumente, W, n_1, n_2, bei uns sind die letzten beid
A_W_p <-  2*(pwilcox(abs(A_W-Erwartungswert)+Erwartungswert, n, n, lower.tail = FALSE) )
B_W_p <-  2*(pwilcox(abs(B_W-Erwartungswert)+Erwartungswert, n, n, lower.tail = FALSE) )
```

Beantworten Sie nun folgende Fragen:

- Wie oft war der *p*-Wert für Stichprobe A kleiner als .05? [ ] α-Fehler
- Wie oft war der *p*-Wert für Stichprobe B kleiner als .05? [ ]
- Wie oft war der *p*-Wert für Stichprobe A grösser als .05? [ ]
- Wie oft war der *p*-Wert für Stichprobe B grösser als .05? [ ] β-Fehler
- Im Vergleich zu unseren Ergebnissen für Konfidenzintervalle und parametrische Signifikanztests, was können Sie schlussfolgern?