7.3 Chi-Quadrat-Streuungstest und F-Test

Alle bisher besprochenen Statistischen Tests sind sog. Tests über die Mittelwerte; denn ihre Nullhypothesen handeln vom Vergleich entweder zweier Mittelwerte oder eines Mittelwertes mit einem konstanten Wert.

In verschiedenen realen Sachverhalten ist es aber sinnvoller, nicht die Mittelwerte zu testen, sondern die Varianzen, da es in der zugrundeliegenden Fragestellung um die Streuung von Messwerten geht. Insbesondere wenn die Streuung recht groß wird, besitzt dann der Mittelwert nur noch wenig Aussagekraft. In solchen Fällen sind Nullhypothesen über Varianzen erheblich sinnvoller. (Allerdings kommen diese Fälle auch seltener vor als Fälle, in denen Tests über die Mittelwerte angesagt sind.)

In diesem Zusammenhang behandeln wir hier zwei Tests: Einmal den Chi-Quadrat-Streuungstest für den Ein-Stichproben-Fall und einmal den F-Test für den Fall zweier voneinander unabhängiger Stichproben.

6.) <u>Chi-Quadrat-Streuungstest</u> (für eine Stichprobe)

Gegeben sei eine Stichprobe $X_1, \ldots, X_n$ i.i.d. $N(\mu; \sigma^2)$- verteilter Zufallsvariablen

mit ______________________________ Erwartungswert μ

und ______________________________ Varianz σ^2.

<u>Fall 1:</u> (zweiseitiger Test)	<u>Fall 2:</u> (einseitiger Test)	<u>Fall 3:</u> (einseitiger Test)
$H_0: \sigma^2 = \sigma_0^2$	$H_0: \sigma^2 \leq \sigma_0^2$	$H_0: \sigma^2 \geq \sigma_0^2$
$H_1: \sigma^2 \neq \sigma_0^2$	$H_1: \sigma^2 > \sigma_0^2$	$H_1: \sigma^2 < \sigma_0^2$

Testgröße: (für alle drei Fälle)

wobei

die Stichprobenvarianz ist und $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ der Stichprobenmittelwert.

Verteilung der Testgröße:

Falls wir annehmen, dass ____________ gilt, dann ist ________________________

____________________________ .

Ablenhungsbereiche:

Ablehnung von H_0, falls

Fall 1: $T(x_1, \ldots, x_n)$ ________________ oder $T(x_1, \ldots, x_n)$ ________________

Fall 2: $T(x_1, \ldots, x_n)$

Fall 3: $T(x_1, \ldots, x_n)$

wobei α das (vorher festgelegte) Signifikanzniveau des Tests ist. Die Quantile der $\chi^2_{n-1}-$Verteilung finden sich in Tabelle C.

Beispiel 51 Ein Möbelhersteller möchte eine neue Regalserie produzieren. Um den richtigen Abstand der Regalböden voneinander planen zu können, besorgt sich der Möbelhersteller die Höhenmaße von ____________ Buchtypen. Aus diesen Maßen errechnet er eine Stichprobenvarianz von ____________________ .

Mit dem Chi-Quadrat-Anpassungstest soll nun überprüft werden, ob die Varianz _____ der Buchhöhen den Wert _________________________ , der beim Bau der alten Regalserie zugrundegelegt wurde, überschreitet oder nicht; d.h. ob bei der neuen Regalserie ein größerer Abstand der Regalböden voneinander eingeplant werden soll als bei der alten Regalserie oder nicht.

Getestet werden soll also die Nullhypothese ___________________________________ gegen die Alternativhypothese ___________________________________ . Wir legen ein Signifikanzniveau von _______________ zugrunde.

Für die Testgröße berechnet sich ein Wert von

Aus Tabelle C haben wir: ____ . Wegen ______________________________________ kann die Nullhypothese selbst auf diesem großzügigen Signifikanzniveau nicht abgelehnt werden, d.h. es ist nicht nötig, bei der neuen Regalserie einen größeren Abstand einzuplanen als bei der alten.

7.) <u>F-Test</u> (für zwei unabhängige Stichproben)

Gegeben seien zwei voneinander unabhängige Stichproben $X_1, \ldots, X_{n_1}$ i.i.d. $N(\mu_1; \sigma_1^2)$ verteilter Zufallsvariablen und $Y_1, \ldots, Y_n$ i.i.d. $N(\mu_2; \sigma_2^2)$-verteilter Zufallsvariablen mit ____________________ Erwartungswerten μ_1, μ_2 und ____________________ Varianzen σ_1^2, σ_2^2.

Fall 1: (zweiseitiger Test)	Fall 2: (einseitiger Test)	Fall 3: (einseitiger Test)
$H_0: \sigma_1^2 = \sigma_2^2$	$H_0: \sigma_1^2 \leq \sigma_2^2$	$H_0: \sigma_1^2 \geq \sigma_2^2$
$H_1: \sigma_1^2 \neq \sigma_2^2$	$H_1: \sigma_1^2 > \sigma_2^2$	$H_1: \sigma_1^2 < \sigma_2^2$

<u>Testgröße:</u> (für alle drei Fälle)

wobei

die Stichprobenvarianz der ersten Stichprobe ist und

die Stichprobenvarianz der Zweiten. Sowie $\bar{X} = \frac{1}{n_1}\sum_{i=1}^{n_1} X_i$ und $\bar{Y} = \frac{1}{n_2}\sum_{i=1}^{n_2} Y_i$.

<u>Verteilung der Testgröße:</u>

Falls wir annehmen, dass __________ gilt, dann ist ______________________________

______________________________ .

Ablenhungsbereiche: Ablehnung von H_0, falls

Fall 1: $T(x_1, \ldots, x_{n_1}, y_1, \ldots, y_{n_2})$ oder $T(x_1, \ldots, x_{n_1}, y_1, \ldots, y_{n_2})$

Fall 2: $T(x_1, \ldots, x_{n_1}, y_1, \ldots, y_{n_2})$

Fall 3: $T(x_1, \ldots, x_{n_1}, y_1, \ldots, y_{n_2})$

wobei α das (vorher festgelegte) Signifikanzniveau des Tests ist. $n_1 - 1$ sind die Zählerfreiheitsgrade und $n_2 - 1$ sind die Nennerfreiheitsgrade. (Die Quantile der F_{n_1-1,n_2-1}–Verteilung finden sich in Tabelle E.)

Hinweis:

Ein F-Test auf Gleichheit der Varianzen (also Fall 1) wird oft einem t-Test für zwei unabhängige Stichproben vorgeschaltet, um zu überprüfen, ob die Voraussetzung gleicher Varianzen in beiden Stichproben abzulehnen ist oder nicht. Statistik-Software-Computer-Programme machen dies z.T. automatisch und geben ggf. eine Warnmeldung heraus, falls die Annahme gleicher Varianzen durch den F-Test nicht bestätigt wird.

Beispiel 52 *Ein Möbelhersteller möchte eine neue Regalserie mit modernem Design herstellen, die eine alte Regalserie mit altmodischem Design ablösen soll. Um den richtigen Abstand der Regalböden voneinander planen zu können, liegen dem Möbelhersteller für die alte Regalserie die Höhenmaße von ____________ damals auf dem Markt befindlichen unterschiedlichen Buchtypen vor mit Stichprobenvarianz ____________ und für die neue Regalserie liegen dem Möbelhersteller*

Höhenmaße von ________________ *heute auf dem Markt befindlichen unterschiedlichen Buchtypen vor mit Stichprobenvarianz* _______________________ .

Beim ersten Durchsehen seiner Daten glaubt der Möberhersteller bei den Höhenmaßen der neuen Buchtypen eine kleinere Varianz zu erkennen als bei den Höhenmaßen der alten Buchtypen; jedoch die Mittelwerte von alten und neuen Höhenmaßen fallen ziemlich gleich aus. Die Fragestellung lautet nun: Kann bei der modernen Regalserie ein kleinerer Abstand der Regalböden voneinander eingeplant werden als bei der alten Regalserie?

Dazu soll mit einem $F-Test$ *zum Signifikanzniveau* ______________ *geprüft werden, ob die Varianz der neuen Buchhöhen kleiner ist als die Varianz der alten Buchhöhen oder nicht.*

Es soll also die Nullhypothese ____________________________ *gegen die Alternativhypothese* ____________________________ *getestet werden (Fall* ____ *). Wir berechnen dazu den Wert der Testgröße aus den beiden Stichproben und erhalten:*

Aus Tabelle E: _____________________________________ .

Da nun gilt: ________________________________ *, fällte der Wert der Testgröße* ______________ *in den Ablehnungsbereich, d.h.* H_0 *kann auf diesem Signifikanzniveau* ______________ *verworfen werden; was* ________________ *spricht, bei der neuen Regalserie einen kleineren Abstand einzuplanen als bei der alten.*

7.4 U-Test von Mann-Whitney und Wilcoxon-Test

In allen bisher behandelten Statistischen Tests wurde stets vorausgesetzt, dass die Zufallsvariablen, welche die Stichprobendaten (oder die Beobachtungsdaten) beschreiben, normalverteilt sind. In vielen Situationen kann man aber nicht davon ausgehen!

Oftmals sind die zugrundeliegenden Stichprobendaten (oder Beobachtungsdaten) noch nicht einmal verhältnis-skaliert, sondern lediglich intervallskaliert oder sogar nur ordinal skaliert.

In der Statistik wurden dafür spezielle Tests entwickelt, die sog. ______________________

________________________________ oder ____

Tests. Wir werden im folgenden zwei solcher Tests vorstellen:

8.) <u>U-Test von Mann-Whitney</u> (für zwei unabhängige Stichproben)

Dieser Test eignet sich für mindestens ordinalskalierte Beobachtungsdaten, welche wir uns reell codiert vorstellen.

Gegeben seien zwei voneinander unabhängige Stichproben $X_1, \ldots, X_n$ i.i.d. Zufallsvariablen gemäß einer Verteilungsfunktion ____________________ , und $Y_1, \ldots, Y_n$ i.i.d. Zufallsvariablen gemäß einer (verschobenen) Verteilungsfunktion

_______________________________________ , für ein ____________ .

Getestet wird dann:

H_0 : ____

H_1 : ____

(Merkhilfe: H_0 : Gleiche Wirkung, H_1 : unterschiedliche Wirkung)

<u>Testgröße:</u>

wobei

Um den Wert der Testgröße zu ermitteln, verfahren wir folgendermaßen: Wir sortieren alle Stichprobenrealisierungen $x_1, \ldots, x_{n_1}, y_1, \ldots, y_{n_2}$ aus den beiden Stichproben der Größe nach und vergeben Rangplätze von _____ für den kleinsten Stichprobenwert bis _______________ für den größten Stichprobenwert. Dann betrachten wir jeden einzelnen Stichprobenwert _____ aus der _____ Stichprobe und zählen, wie viele Stichprobenwerte _____ aus der _____ Stichprobe einen echt größeren Rangplatz haben als _____ .

Dies ergibt die Anzahl _________________ der sog. Rangplatzüberschreitungen für jedes einzelne _____ aus der _____ Stichprobe. Alle diese Rangplatzüberschreitungen aufsummiert ergibt den Wert der Testgröße, nämlich:

Dazu ein Beispiel: Es liegen uns folgende Stichprobenrealisierungen aus den beiden Stichproben vor, denen wir ihrer Größe nach Rangplätze zuweisen. In der letzten Zeile schließlich stehen bei jedem Stichprobenwert der ersten Stichprobe die Anzahl der Stichprobenwerte aus der zweiten Stichprobe, die einen größeren Rangplatz aufweisen:

Diese Zahlen der letzten Zeile aufsummiert ergeben den Wert der Testgröße:

____

Verteilung der Testgröße:

Falls wir annehmen, dass __________ gilt, und falls entweder _______________ oder _______________ (also bei mindestens einer hinreichend großen Stichprobe),

dann ist ____ näherungsweise
____ mit ___________________________

und ____ .

Ablenhungsbereich: Ablehnung von H_0, falls

wobei α das (vorher festgelegte) Signifikanzniveau des Tests ist und das $z_{1-\alpha/2}-$Quantil der Standardnormalverteilung in der z-Tabelle zu finden ist.

Vorliegen von Bindungen:

Je nachdem, wie die beiden Stichproben ausfallen, kann es vorkommen, dass unter den Stichprobenrealisierungen $x_1, \ldots, x_{n_1}, y_1, \ldots, y_{n_2}$ Werte ____________________ vorkommen, d.h. die Rangplätze können nicht mehr ____________________ vergeben werden. In diesem Fall spricht man vom Vorliegen von sog. Bindungen und wir verfahren wie folgt:

Wir sortieren alle Stichprobenrealisierungen $x_1, \ldots, x_{n_1}, y_1, \ldots, y_{n_2}$ der Größe nach, nur dass jetzt die mehrfach vorkommenden Werte nebeneinander stehen. Dann vergeben wir Rangplätze an die einzeln vorkommenden Werte ____________________ , und an die mehrfach vorkommenden Werte wird jedem der ____________________ der für diese Werte normalerweise zu vergebenden Rangplätze zugewiesen.

Dazu ein Beispiel: Seien 1; 1; 2; 3; 3; 3; 4 die bereits der Größe nach sortierten Stichprobenrealisierungen aus den beiden Stichproben. Dann werden ihnen folgende Ränge zugeordnet:

Die Testgröße beim Vorliegen von Bindungen ist wie bisher die Summe der Rangplatzüberschreitungen:

wobei wieder

Verteilung der Testgröße:

Falls wir annehmen, dass ________ gilt, und falls entweder $n_1 > 10$ oder $n_2 > 10$ (also bei mindestens einer hinreichend großen Stichprobe), dann ist $U(X_1, \ldots, X_{n_1}, Y_1, \ldots, Y_{n_2})$ näherungsweise ______________ -verteilt mit ______________________ und

wobei ____ = Anzahl der verschiedenen Werte, die jeweils mehrfach vorkommen, und zwar mit Häufigkeiten ________________ .

Zu unserem letzten Beispiel:

Hier ist __________ und ______________________________ (Erklärung: Es gibt zwei verschiedene Stichprobenwerte, die mehrfach vorkommen, nämlich die _____ und die _____ . Also ist __________ . _____ kommt zweimal vor, also ist ____________ . _____ kommt dreimal vor, also ist ____________).

Für die Summe berechnet sich damit:

Ablehnungsbereich beim Vorliegen von Bindungen:

Ablehnung von H_0, falls

wobei α wieder das Signifikanzniveau des Tests ist und das $z_{1-\alpha/2}-$Quantil in der z-Tabelle zu finden ist.

Bemerkung: In beiden Fällen (sowohl ohne als auch mit Vorliegen von Bindungen) kann man sagen, dass die Testgröße den Grad der Durchmischung der Stichprobenwerte beider Stichproben misst. Je schlechter durchmischt die Werte beider Stichproben sind, desto extremer wird der Wert der Testgröße ausfallen, d.h. die Abweichung von μ_u wird entweder in die eine oder in die andere Richtung sehr groß, und desto eher wird es zu einer Verwerfung der Nullhypothese kommen. Je besser durchmischt die Werte beider Stichproben sind, desto mehr nähert sich der Wert der Testgröße μ_u an, und desto unwahrscheinlicher wird eine Verwerfung der Nullhypothese.

Beispiel 53

(In den Stichproben zu diesem Beispiel kommen keine Bindungen vor; diesen Fall werden wir in den Übungen behandeln.)

Ein Arzt gibt 12 Patienten, die unter Schlafstörungen leiden, für eine Nacht Medikament A und 14 anderen Patienten, die ebenfalls unter Schlafstörungen leiden, gibt er für eine Nacht Medikament B.

Medikament A		*Medikament B*	
Schlafdauer	*Rangplatz*	*Schlafdauer*	*Rangplatz*
3 : 55	8	9 : 25	26
4 : 08	10	3 : 37	6
8 : 11	23	5 : 09	13
2 : 46	5	1 : 18	1
1 : 43	2	2 : 25	3
7 : 23	21	4 : 53	12
6 : 14	17	3 : 59	9
9 : 01	25	7 : 13	20
5 : 13	14	4 : 18	11
8 : 33	24	6 : 45	19
7 : 29	22	3 : 48	7
5 : 38	16	2 : 37	4
		6 : 41	18
		5 : 27	15

Der Arzt vermutet eine unterschiedliche Wirkungsweise beider Medikamente und möchte dies mit Hilfe eines U-Tests von Mann-Whitney untersuchen. Dazu lässt er sich von allen Patienten aufschreiben, wielange sie in der einen Nacht geschlafen haben. Es liegen ihm folgende Werte (mit zugehörigen Rangplätzen) in Std.:Min. vor.

Wir müssen nun den Wert der Testgröße, d.h. die Summe der Rangplatzüberschreitungen, ermitteln. Dazu betrachten wir jeden einzelnen Wert aus dem linken Tabellenteil, also jede einzelne Schlafenszeit unter dem Einfluss von Medikament A, merken uns ihren Rangplatz, und zählen, wieviele Werte aus dem rechten Tabellenteil (unter Medikament B) einen größeren Rangplatz haben.
Für den ersten Wert ____________ *mit Rangplatz* _____ *haben z.B.* _____ *Werte unter Medikament B einen größeren Rangplatz, für den zweiten Wert* ____________ *mit Rangplatz* _______ *haben* _____ *Werte unter B einen größeren Rangplatz, usw..*
Alle diese Rangplatzüberschreitungen aufsummiert ergibt den Wert der Testgröße:

Wegen ___________________________________ *kann bei Anwendung des U-Tests von Mann-Whitney auf diese Testsituation die Nullhypothese* _________________ *verworfen werden, d.h. die Daten geben keinen Anlass zur Vermutung des Arztes, dass die Medikamente* ______________________________________ *Wirkung haben.*

9.) Wilcoxon-Test (oder auch: Vorzeichen-Rang-Test)

(für zwei abhängige Stichproben)

Dieser Test eignet sich für mindestens intervallskalierte Beobachtungsdaten.

Gegeben seien zwei abhängige Stichproben i.i.d. Zufallsvariablen $X_1, \ldots, X_n$ (das sind die Stichprobenergebnisse ____________________) und i.i.d. Zufallsvariablen $Y_1, \ldots, Y_n$ (das sind die Stichprobenergebnisse ____________________).

Dann sind auch die Differenzen _________________________ für _______________ i.i.d. Zufallsvariablen. (Merkhilfe: Die Differenzen kann man sich wie die Veränderungen von vorher zu nachher vorstellen.)

Getestet wird dann:

H_0 : ____ für alle $x \geq 0$, $i = 1, \ldots, n$, d.h. die Differenzen $D_1, \ldots, D_n$ (also die "Veränderungen") sind symmetrisch um Null verteilt

H_1 : Die Differenzen $D_1, \ldots, D_n$ sind nicht symmetrisch um Null verteilt, d.h. es gibt ein $x \geq 0$ und ein $i \in \{1, ..., n\}$ mit ____ .

(Merkhilfe: H_0 : Keine Veränderung zwischen vorher und nachher,

H_1 : Es hat sich etwas verändert)

Testgröße:

wobei _____ der Rang vom Absolutbetrag der Differenz _______ ist. Aufsummiert werden nur diejenigen Ränge, die zu echt positiven Differenzen gehören.

Den Wert der Testgröße ermitteln wir folgendermaßen:

Wir bilden die Differenzen ____________________________________ , und ordnen die Absolutbeträge ________ der Differenzen der Größe nach. Dann vergeben wir Rangplätze von _____ für den kleinsten Absolutbetrag bis _____ für den größten Absolutbetrag. Schließlich addieren wir diejenigen Ränge auf, die zu echt positiven Differenzen _____ gehören, für die also __________ gilt. Diese Summe ergibt den Wert der Testgröße.

Verteilung der Testgröße:

Falls wir annehmen, dass H_0 gilt, und falls $n \geq 20$ (also bei großen Stichproben), dann ist ____ näherungsweise ____ mit ____________________________ und ____________________________________ .

Ablenhungsbereich:

Ablehnung von H_0, falls

wobei α das Signifikanzniveau des Tests ist und das $z_{1-\alpha/2}-$Quantil der Standardnormalverteilung in der z-Tabelle zu finden ist.

Vorliegen von Bindungen:

Je nachdem, wie die beiden Stichproben ausfallen, kann es vorkommen, dass unter den Differenzen ______________________________ Werte mehrfach vorkommen, d.h. die Rangplätze können nicht mehr eindeutig vergeben werden. Ebenso kann es vorkommen, dass ein oder mehrere Differenzen gleich Null sind. In diesen Fällen spricht man vom Vorliegen sog. "Bindungen" und wir verfahren wie folgt:

Wir sortieren die Absolutbeträge der Differenzen ____________________ wie bisher der Größe nach, nur dass jetzt die mehrfach vorkommenden Werte nebeneinander stehen und auch Nullen (einfach oder mehrfach) vorkommen können. Dann vergeben wir Rangplätze an die einzelnen Werte wie gewohnt. Dabei werden die Nulldifferenzen als "kleinste" Differenzen berücksichtigt. An die mehrfachen Werte wird jedem der Durchschnitt der für diese Werte normalerweise zu vergebenden Rangplätze zugewiesen.

Beispiel dazu: Sei n=8 und

$$y_1 = 1; \quad y_2 = 2; \quad y_3 = 2; \quad y_4 = 3; \quad y_5 = 0; \quad y_6 = 2; \quad y_7 = 3; \quad y_8 = -1$$

$$x_1 = 2; \quad x_2 = 0; \quad x_3 = 2; \quad x_4 = 1; \quad x_5 = 0; \quad x_6 = 1; \quad x_7 = 1; \quad x_8 = 3$$

Jedoch gehören nur Rangplätze mit Pfeilen zu echt positiven Differenzen (die anderen Rangplätze gehören zu negativen Differenzen oder zu Null-Differenzen).

Die Testgröße ist wie bisher die Summe derjeniger Rangplätze, die zu echt positiven Differenzen gehören; hier: ______________________________ .

Also:

wobei ______ der Rang des Absolutbetrages der Differenz ______ ist.

<u>Verteilung der Testgröße:</u> Falls wir annehmen, dass H_0 gilt, und falls $n \geq 20$ (also bei großen Stichproben), dann ist ______________________________

näherungsweise ______________________________ mit

und

wobei k = Anzahl der verschiedenen Differenzen-Absolutbeträge $\neq 0$, die jeweils mehrfach vorkommen, und zwar mit Häufigkeiten ______________ ;

t_0 = Häufigkeit der vorkommenden Null-Differenzen ________________ .

Zu unserem letzten Beispiel:

Hier ist __________ (weil ____ Nulldifferenzen vorkommen) und __________ (weil ____ verschiedene Absolutbeträge $\neq 0$ mehrfach vorkommen, nämlich die ____ mit Häufigkeit ____________ und die ____ mit Häufigkeit ____________).

Somit ist

Ablehnungsbereich bei Vorliegen von Bindungen:

Ablehnung von H_0, falls

wobei α wieder das Signifikanzniveau des Tests ist und das $z_{1-\alpha/2}$-Quantil aus der z-Tabelle.

Beispiel 54 *(Zum letzten Beispiel)*

Als "vorher-nachher-Problem" formuliert:

Ein Arzt gibt 20 Patienten, die unter Schlafstörungen leiden, für eine Nacht Medikament A, und lässt sich von allen Patienten aufschreiben, wie lange sie in der einen Nacht (unter Einwirkung von Medikament A) geschlafen haben.

Zusätzlich lässt er sich von allen Patienten berichten, wie lange sie in der Nacht davor (ohne Medikament A) geschlafen haben. Der Arzt vermutet eine Wirkung von Medikament A, und möchte dies mit einem Wilcoxon-Test untersuchen. Es liegen ihm folgende Werte vor (in Std. : Min.), siehe Tabelle auf der nächsten Seite.

$x_i =$ ____

$y_i =$ ____

$r_i =$ __________________

x_i	y_i	$d_i = y_i - x_i$	$\lvert d_i \rvert$	r_i
3:55	5:02	1:07	1:07	8
4:08	6:27	2:19	2:19	14
2:46	8:24	5:38	5:38	20
8:11	5:03	-3:08	3:08	18
7:23	6:25	-0.58	0.58	7
1:43	2:58	1:15	1:15	10
9:01	6:31	-2.30	2:30	16
5:13	8:37	3:24	3:24	19
5:38	6:15	0:37	0:37	3
7:29	6:45	-0:44	0:44	4
8:33	8:18	-0:15	0:15	1
3:37	6:38	3:01	3:01	17
5:09	6:51	1:42	1:42	12
1:18	2:13	0:55	0:55	6
7:13	6:21	-0:52	0:52	5
2:25	4:38	2:13	2:13	13
9:25	6:59	-2:26	2.26	15
4:53	6:03	1:10	1:10	9
3:59	5:20	1:21	1:21	11
4:15	4:47	0:32	0:32	2

Die Testgröße ______________________________ *ist nun die Summe aller Rangplätze* _____ *, die zu positiven Differenzen* ____________ *gehören, also:* ______________________________________ *.*

Für den Ablehnungsbereich müssen wir noch _______ *und* _______ *berechnen (für* $n = 20$*):* ______________________________ *und* ______________________________________ *.*

Damit:

Aus der z-Tabelle (für $\alpha = 0,1$*):*

Da ______________________________ *, kann* H_0 *auf diesem (schon recht großzügigen) Signifikanzniveau* ______________________________ *werden, was bedeutet, dass eine Auswirkung von Medikament A auf die Schlafenszeit der Patienten* ___________ *nachgewiesen werden kann.*

7.5 Kolmogorow-Smirnow-Test, Chi-Quadrat-Anpassungstest und Chi-Quadrat-Unabhängigkeitstest

Die Statistischen Tests, welche wir in den Abschnitten 7.2 und 7.3 behandelt haben, benötigen als Voraussetzung, dass die Stichprobendaten einer normalverteilten Grundgesamtheit angehören.

In vielen Fällen kann man aus der Erfahrung heraus sagen, dass diese Voraussetzung gegeben ist. In manchen Fällen jedoch bestehen Zweifel, und in diesen Fällen gibt es die Möglichkeit, zuerst einen sog. ______________________________ durchzuführen. Ein ______________________________ oder auch ______________ ______________________________ (= ______________________________ ______________________________) dient zur Überprüfung der Nullhypothese, ob das beobachtete Merkmal eine bestimmte Verteilung, z.B. die Normalverteilung, besitzt oder nicht; d.h. ob sich die Stichprobendaten hinreichend gut an eine gewünschte Verteilung "anpassen".

Wir werden hier zwei solcher Tests kennenlernen:

Den ____ und

den ____________________________________ .

10.) <u>Kolmogorow-Smirnow-Test</u> (für eine Stichprobe)

Gegeben sei eine Stichprobe i.i.d. Zufallsvariablen$X_1, \ldots, X_n$ aus einer Grundgesamtheit mit einer ____________________________ stetigen Verteilungsfunktion F. Ferner sei F_0 eine ________________________ stetige Verteilungsfunktion, z.B. der Normalverteilung $N(\mu; \sigma^2)$ mit bekannten Erwartungswert μ und bekannter Varianz σ^2.

<u>Testgröße:</u>

Für die Testgröße berechnen wir zuerst die empirische Verteilungsfunktion F_n aus den Stichprobenrealisierungen $x_1, \ldots, x_n$ wie folgt:

mit $m(x)$ gleich der Anzahl der Stichprobenrealisierungen, die x nicht übertreffen.

Beispiel: Gegeben sei eine Stichprobe vom Umfang n=6 wie folgt:

$$x_1 = 2.5; \qquad x_2 = -1; \qquad x_3 = 1; \qquad x_4 = 2.5; \qquad x_5 = 2; \qquad x_6 = 3$$

Diese Stichprobe besitzt folgende empirische Verteilungsfunktion:

Sei nun F_0 die Verteilungsfunktion der bekannten $N(1.5; 2)$-Verteilung. Wir betrachten die Graphen der empirischen Verteilungsfunktion F_6 (= durchgezogene Linie) und der bekannten Verteilungsfunktion F_0(= gestrichelte Linie):

Die **Testgröße** ist nun der maximale Abstand von bekannter Verteilungsfunktion F_0 und empirischer Verteilungsfunktion F_n. (Dieser wird immer an den Sprungstellen von F_n angenommen. Also in $x_{(i)}, i = 1, \ldots n$.) In Formeln:

<u>Verteilung der Testgröße:</u>

Falls wir annehmen, dass __________ gilt, dann ist $T(X_1, \ldots, X_n)$ verteilt gemäß der sog. ____ -Verteilung. (Für n > 40 können die Quantile der asymptotischen Verteilung genommen werden, siehe Tabelle M.)

<u>Ablehnungsbereiche:</u> Ablehnung von H_0, falls

wobei $k_{1-\alpha}$ das $(1-\alpha)$-Quantil der Kolmogorow-Smirnow-Verteilung ist.

Weitaus häufiger als der Kolmogorow-Smirnow-Test wird der folgende Anpassungstest verwendet:

11.) **Chi-Quadrat-Anpassungstest** **(für eine Stichprobe)**

Dieser Test eignet sich bereits für diskrete Verteilungen; kann aber auch bei stetigen Verteilungen angewendet werden.

Gegeben sei eine Stichprobe i.i.d. Zufallsvariablen $X_1, \dots, X_n$ aus einer Grundgesamtheit mit einer ______________________________ Verteilungsfunktion F (F kann stetig oder diskret sein, beides ist möglich).

Ferner sei F_0 eine ____________________ (diskrete oder stetige) Verteilungsfunktion.

Testgröße:

Zuerst wird die reelle Achse in k sich ausschließende Intervalle eingeteilt (wobei $k \leq n$), und dann wird gezählt, wieviele Stichprobenrealisierungen $x_1, \dots, x_n$ in jedes Intervall fallen; diese Anzahlen werden mit $n_1, \dots, n_k$ bezeichnet.

Sei nun p_i die Wahrscheinlichkeit dafür, dass eine Zufallsvariable Y, welche gemäß der Verteilungsfunktion F_0 verteilt ist, einen Wert in der i-ten Klasse annimmt, d.h.

____ . Die Testgröße ist:

Dabei kann man ________ auch interpretieren als die erwartete Anzahl von Stichprobenrealisierungen im ________ Intervall für $i = 1, \dots, k$.

Verteilung der Testgröße:

Falls wir annehmen, dass ________ gilt, dann ist $T(X_1, \ldots, X_n)$ näherungsweise

______________________________ .

Ablehnungsbereich:

Ablehnung von ____ , falls

wobei $\chi^2_{k-1;1-\alpha}$ das $(1-\alpha$-Quantil der χ^2-Verteilung mit $k-1$ Freiheitsgraden ist und in Tabelle C abgelesen werden kann.

Beispiel 55

Mit Hilfe des χ^2-Anpassungstests soll untersucht werden, ob ein vorliegender Würfel "fair" oder "verfälscht" ist. Dazu wird der Würfel 100 mal geworfen und die geworfenen Augenzahlen notiert:

Augenzahl	*1*	*2*	*3*	*4*	*5*	*6*
Anzahl der zugehörigen Würfe	*20*	*17*	*15*	*19*	*13*	*16*

Wir haben also $k = 6$ Klassen vorliegen . Wir testen mit folgender Nullhypothese:

Das heißt:

(Jede Augenzahl bei einem fairen Würfel kommt mit der Wahrscheinlichkeit $\frac{1}{6}$ vor.)

Damit ergibt sich als Wert der Testgröße:

Für ________________ *lesen wir in Tabelle C ab:*

d.h. ____ *, und*

damit kann H_0 ______________ *verworfen werden, d.h. die Daten sprechen* ______________ *gegen einen fairen Würfel!*

<u>Hinweis:</u>

Im Gegensatz zu allen Tests, die wir in 7.2 bis 7.4 vorgestellt haben, steht in den beiden hier vorgestellten Anpassungstests die erwünschte Situation in der ____________________________________, nicht in der ____! D.h. hier ist es günstig, wenn H_0 ______________ verworfen werden kann!

Zuguterletzt wollen wir noch den sog. ______________________________

______________ behandeln, der die Nullhypothese testet, ob zwei Merkmale in derselben Grundgesamtheit ______________________ voneinander sind.

Anwendungen dieses Tests finden sich in vielen Fragestellungen, z.B.:

Beispiel 56

(a) *Gibt es einen Zusammenhang zwischen Körpergröße und Körpergewicht gleichalter Erwachsener?*

(b) *Sind Augenfarbe und Haarfarbe voneinander unabhängige Merkmale?*

(c) *Beeinflusst das Einkommen von Wahlberechtigten ihre Wahlentscheidung?*

(d) *Ist die Blutgruppenzugehörigkeit geschlechtsabhängig?*

12.) Chi-Quadrat-Unabhängigkeitstest (für zwei abhängige Stichproben)

Gegeben seien zwei voneinander ___________________ Stichproben

$X_1, \ldots, X_n$ i.i.d. Zufallsvariablen des Merkmals X und

$Y_1, \ldots, Y_n$ i.i.d.Zufallsvariablen des Merkmals Y.

Testgröße:

Beide reellen Achsen, sowohl die x-Achse als auch die y-Achse, werden in sich gegenseitig ausschließende Intervalle eingeteilt, und zwar die x-Achse in __________ Intervalle und die y-Achse in __________ Intervalle. Dann wird gezählt, wieviele Stichprobenrealisierungen $x_1, \ldots, x_n$ von Merkmal X in die k Intervalle der x-Achse fallen und wieviele Stichprobenrealisierungen $y_1, \ldots, y_n$ von Merkmal Y in die l Intervalle der y-Achse fallen. Diese Anzahlen werden in eine Kontingenztafel eingetragen und wie folgt bezeichnet:

Mit diesen Bezeichnungen aus der Kontingenztafel ist die Testgröße:

Dabei kann man ____________ interpretieren als die erwartete Häufigkeit in Intervall i von Merkmal X und (gleichzeitig) in Intervall j von Merkmal Y (unter H_0).

Verteilung der Testgröße:

Falls wir annehmen, dass _____ und _____ ____________________________ sind, dann ist die Testgröße näherungsweise _________________________ -verteilt.

Ablehnungsbereiche:

Ablehnung von H_0, falls

wobei α das Signifikanzniveau des Tests ist und ______________________________ das ______________________________ der ____ .

Anmerkung:

Der χ^2 -Unabhängigkeitstest ist geeignet für __________ Skalenart der Merkmale X und Y.

Beispiel 57 *(Zu Teil (c) des letzten Beispiels)*

Es soll untersucht werden, ob ein Zusammenhang zwischen Einkommen und Wählerverhalten besteht. Dazu werden 1000 *zufällig ausgewählte Bundesbürger nach ihrem Einkommen (Ausprägungen: hoch, mittel, niedrigi) und der Partei (Ausprägungen: A, B, C, andere) befragt, der sie bei der nächsten Bundestagswahl ihre Stimme geben wollen. Es ergibt sich folgende Kontingenztafel für die absoluten Häufigkeiten:*

Für die Testgröße wurde noch in jedem Feld der Kontingenztafel, d.h. für jedes Indexpaar $i \in \{1, \ldots, k\}$ und $j \in \{1, \ldots, l\}$ die Werte ______________ *berechnet.*

Mit diesen Werten berechnet sich die Testgröße wie folgt:

Für ein Signifikanzniveau von $\alpha = 0.05$ *lesen wir in Tabelle C ab:*

Wegen ____ *lehnen wir die Nullhypothese* H_0 *: "Einkommen und Wahlverhalten sind unabhängig" ab, d.h. wir haben mit dem* χ^2*-Unabhängigkeitstest einen Zusammenhang zwischen Einkommen und Wahlverhalten nachgewiesen.*

Übersicht Statistische Tests OHNE Normalverteilungsannahme: