iprom

INSTITUT FÜR PRODUKTIONSMESSTECHNIK
TECHNISCHE UNIVERSITÄT BRAUNSCHWEIG

Klausur

Statistische Messdatenauswertung für Biotechnologen

im Modul

BP 14 Statistik und Programmieren

26. März 2024

Name: ______________________________

Matrikel-Nr.: ______________________________

Prüfungsraum: ______________________________

Mit meiner Unterschrift versichere ich, dass ich mich geistig und körperlich in der Lage befinde, die Prüfung abzulegen (d. h. prüffähig bin).

____
Unterschrift Studierende/r

AUFGABE	1	2	AWV A	AWV B	KF	Gesamt
PUNKTE	/17	/11	/18	/17	/6	/69

NOTE

Hinweise zur Prüfung

1. Bearbeitungsdauer: 120 Minuten
2. Als Hilfsmittel sind ausschließlich Taschenrechner ohne vorgefertigte Programme und ohne drahtlose Kommunikationsschnittstelle, einschließlich deren Bedienungsanleitung in gedruckter Form, zugelassen. Sonstige schriftliche Unterlagen sowie Bild-, Ton- und Videodokumente sind ausdrücklich nicht zugelassen. Die Verwendung elektronischer Geräte mit drahtloser Kommunikationsschnittstelle, gleich zu welchem Zweck, ist während der Klausur untersagt. Dieses Verbot gilt insbesondere auch für sogenannte Smartwatches. Nach allgemeinem Prüfungsrecht und aktueller APO stellt bereits das Mitführen eines nicht erlaubten Hilfsmittels im Prüfungsraum eine Täuschung dar. Verstöße dagegen bzw. andere Täuschungsversuche werden gemäß der Prüfungsordnung geahndet.
3. Auf das Deckblatt sind der Name, der Vorname, die Matrikelnummer und die Bezeichnung des Raumes, in welchem die Prüfung abgelegt wird einzutragen. Auf allen anderen abgegebenen Blättern ist zumindest der Name zu vermerken. Das Deckblatt ist als oberes Blatt der Klausur abzugeben. Der Rest der Aufgabestellung muss nicht abgegeben werden, sofern er keine für die Lösung relevanten Eintragungen enthält.
4. Zur Teilnahme an der Prüfung ist auf dem Deckblatt die Prüfungsfähigkeit durch Unterschrift zu bestätigen.
5. Der Studierendenausweis ist zusammen mit einem Lichtbildausweis und dem ausgefüllten Deckblatt der Aufgabenstellung sichtbar auszulegen.
6. Alle zur Lösung der gestellten Aufgaben benötigten nichttrivialen Gleichungen und Konstanten sowie alle notwendigen Tabellen und Diagramme sind der folgenden Formelsammlung, der Aufgabenstellung selbst oder dem Anhang auf den Seiten 13 bis 17 zu entnehmen.

Formelsammlung:

Produktregel: $(uv)' = u'v + uv'$

Quotientenregel: $\left(\frac{u}{v}\right)' = \frac{vu' - uv'}{v^2}$

Kettenregel: $\frac{dy}{dx} = u'(v)v'(x) \quad \text{mit} \quad y = u\big(v(x)\big)$

Fakultät: $n! = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot n = \prod_{k=1}^{n} k$

Eulersche Zahl: $\mathrm{e} = 2{,}718281828459045235\ldots$

Elektrische Kapazität: $1\ \mathrm{F}\ (\mathrm{Farad}) = 1\ \mathrm{C/V} = 1\ \mathrm{A \cdot s/V} = 1\ \mathrm{A^2 \cdot s^4/(kg \cdot m^2)}$

1. Aufgabe:

Bei der kapazitiven Füllstandsmessung nutzt man die Tatsache aus, dass Flüssigkeiten eine relative Dielektrizitätskonstante ε_r größer als eins besitzen. Als Messgrößenumformer kommt daher ein Kondensator zum Einsatz, der durch die Behälterwand und eine Stabelektrode gebildet wird (siehe Abbildung 1.1).

Wird der dargestellte Behälter mit Flüssigkeit befüllt, so nimmt gleichzeitig die Kapazität des Kondensators zu. Es existiert ein eindeutiger Zusammenhang zwischen der Gesamtkapazität C_G und der Füllhöhe h, welcher als Berechnungsgrundlage für die Auswertung herangezogen werden kann.

Die zu bestimmende Füllhöhe h ergibt sich gemäß folgendem formelmäßigen Zusammenhang:

$$h = \frac{H}{\varepsilon_r - 1} \cdot \left(\frac{C_G - C_K}{C_E} - 1 \right)$$

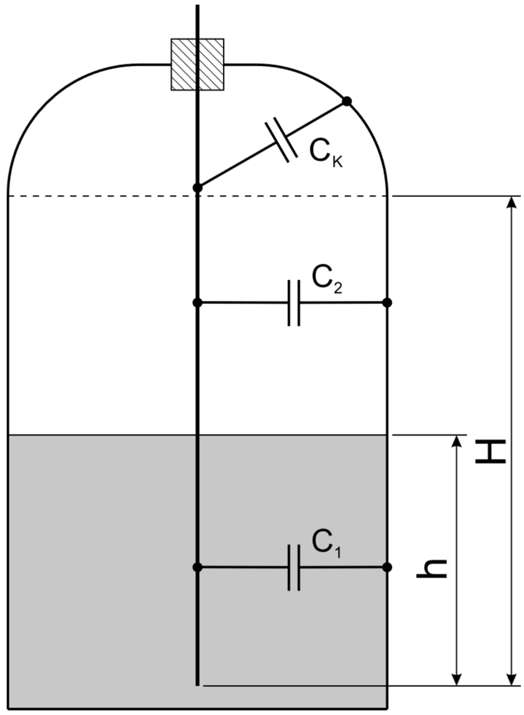


Abbildung 1.1: Prinzip der kapazitiven Füllstandsmessung eines Behälters

Im Folgenden soll die Füllhöhe h des Behälters auf der Grundlage von Messergebnissen für die Größen H, ε_r, C_G, C_K und C_E einschließlich der wahrscheinlichen Abweichungsgrenzen ermittelt werden.

Die maximale Füllhöhe H wird vom Hersteller des Behälters mit einem Wert von $H = 5000\text{ mm} \pm 2{,}5\text{ mm}$ bei $P = 95\%$ und sehr großen Stichprobenumfang n_H angegeben.

Als Füllmedium des Behälters wird Wasser eingesetzt, welches unter den herrschenden Umgebungsbedingungen eine relative Dielektrizitätskonstante von $\varepsilon_r = 80$ aufweist. Dieser Wert kann als exakt angesehen werden.

Die Kapazität C_K wird vom Hersteller des Behälters mit $C_K = 40\text{ pF}$ angegeben. Dieser Wert kann als exakt angesehen werden.

Die Kapazität C_E wurden im Vorfeld experimentell in $n_{C_E} = 25$ Versuchen bestimmt. Das ermittelte Messergebnis beträgt $C_E = 240\text{ pF} \pm 1\text{ pF}$ bei $P = 99\%$.

Die Gesamtkapazität C_G wird im Zuge der Versuchsdurchführung mit insgesamt $n_{C_G} = 8$ Einzelmessungen ermittelt. Die dabei erhaltenen Messwerte sind in Tabelle 1.1 zusammengefasst.

i	1	2	3	4	5	6	7	8
C_G / pF	3405	3362	3381	3417	3426	3346	3347	3329

Tabelle 1.1: Messwerte der Gesamtkapazität C_G

a) Berechnen Sie die gesuchte Füllhöhe h und geben Sie das vollständige Messergebnis mit einer Aussagewahrscheinlichkeit von $P = 95\%$ an!

Hinweis: Für alle Messgrößen kann eine Normalverteilung vorausgesetzt werden.

2. Aufgabe:

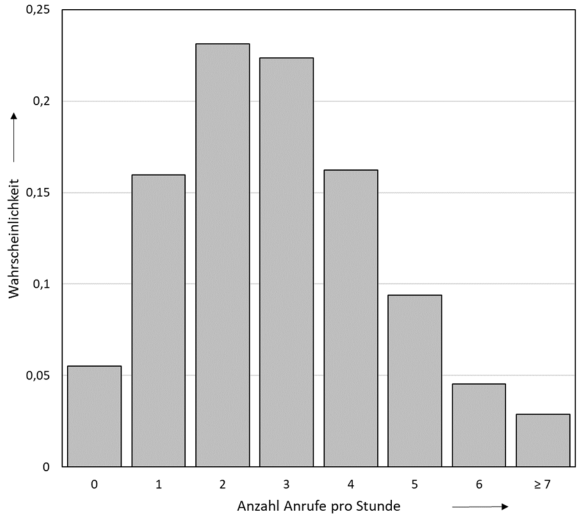


Abbildung 2.1: Poisson-Verteilung mit Parameter $\lambda = 2{,}9$

Bei Ihrem Arbeitgeber, einem Hersteller von Bioreaktoren für die Pharmaindustrie, wird seit einigen Monaten eine Service-Hotline betrieben, bei welcher die Betreiber Störungen ihrer Anlagen melden können, um zeitnah Unterstützung bei der Problembehebung zu erhalten.

Um die Kapazitäten der Service-Hotline und des Service-Teams zukünftig besser planen zu können, analysieren Sie im Rahmen Ihrer Bachelorarbeit die Anzahl der pro Stunde eingehenden Service-Anfragen. Hierzu greifen Sie auf die Datenbank der Serviceabteilung zu und bereiten die Daten derart auf, dass Sie für insgesamt $n = 1800$ Stunden, in welchen die Service-Hotline in Betrieb war, die Anzahl der jeweils verzeichneten Service-Anfragen angeben können. In Tabelle 2.1 ist zusammengefasst, in wie vielen dieser n Stunden jeweils 0, 1, 2, 3, 4, 5, 6 oder 7 und mehr Anfragen eingingen.

Anzahl Anfragen	0	1	2	3	4	5	6	7 und mehr
Häufigkeit	94	296	412	409	281	162	89	57

Tabelle 2.1: Ermittelte Häufigkeiten für Anzahl k der pro Stunde eingegangenen Service-Anfragen

Da die Anzahl der Anfragen gemessen an der Anzahl der in Betrieb befindlichen Anlagen sehr gering ist, haben Sie die Vermutung, dass die beobachtete Verteilung einer Poisson-Verteilung genügt.

Die Wahrscheinlichkeit P_λ dafür, dass bei einem Poisson-verteilten Prozess k Ereignisse registriert werden, ist durch folgenden Ausdruck definiert:

$$P_\lambda(X = k) = \frac{\lambda^k}{k!} \mathrm{e}^{-\lambda}$$

Hierin steht k für die Anzahl der registrierten Ereignisse (im vorliegenden Fall also für die Zahl der pro Stunde eingegangenen Service-Anfragen), λ ist der Parameter der Poisson-Verteilung und e ist die Eulersche Zahl (Basis der natürlichen Exponentialfunktion).

Den Parameter λ haben Sie für die vorliegende Untersuchung anhand des für die Poisson-Verteilung geltenden Zusammenhangs $\mu = \lambda$ im Vorfeld anhand der Tabelle 2.1 zugrundeliegenden Rohdaten zu $\lambda = 2{,}9$ abgeschätzt.

a) Untersuchen Sie mittels eines Chi-Quadrat-Tests, ob die in Tabelle 2.1 angegebene Verteilung auf einem Signifikanzniveau von $\alpha = 0{,}05$ einer Poisson-Verteilung mit dem geschätzten Parameter $\lambda = 2{,}9$ genügt!

Erläuterungen zu Aufgaben nach dem Antwort-Wahl-Verfahren:

Bei jeder Fragestellung wird im Anschluss an die Antwortalternativen angegeben, um welchen Fragetyp es sich handelt. Die möglichen Fragetypen sind nachfolgend näher erläutert.

- *Fragetyp Einfachwahl*: Bei Fragen dieses Typs ist genau eine der angebotenen Antwortalternativen korrekt. Bei Fragen dieses Typs wird nur dann eine von null Punkten verschiedene Bewertung vergeben, wenn genau die eine korrekte Antwort markiert wurde.
- *Fragetyp Mehrfachwahl*: Bei Fragen dieses Typs ist mindestens eine der angebotenen Antwortalternativen korrekt. Entsprechend können auch mehrere oder alle Antwortalternativen korrekt sein. Bei Fragen dieses Typs werden auch dann anteilig Punkte vergeben, wenn einzelne Antworten unzutreffend sind (korrekte Antwort fälschlich nicht markiert oder unkorrekte Antwort fälschlich markiert). Hierbei gilt jedoch, dass eine Frage, bei welcher keine der Antworten markiert wurde als nicht bearbeitet gilt und mit null Punkten bewertet wird.

Für alle Fragetypen gilt, dass eine Frage nicht mit weniger als null Punkten bewertet werden kann. Es werden also keine negativen Punkte vergeben.

Antwort-Wahl-Verfahren, Teil A:

3. Im Rahmen einer routinemäßigen Wareneingangsprüfung soll anhand einer Stichprobe sichergestellt werden, dass die Hall-Konstante A_H gelieferter Sensorplättchen der vereinbarten Spezifikation entspricht. Hierzu wird aus der laufenden Fertigung eine Stichprobe vom Umfang $n = 15$ entnommen und die Hall-Konstante A_H der Sensorplättchen ermittelt. Aus der Stichprobe ergibt sich ein Mittelwert der Hall-Konstanten von $\overline{A_H} = 5{,}02 \cdot 10^{-3}\ \mathrm{m^3/C}$ und eine Streuung von $S_{A_H} = 8{,}4 \cdot 10^{-5}\ \mathrm{m^3/C}$. Die Standardabweichung σ sei unbekannt.

3.1. Das Konfidenzintervall des Erwartungswertes der Hall-Konstanten A_H für eine Aussagewahrscheinlichkeit von P = 99% beträgt für diesen Fall ungefähr:

a) $A_H = 5{,}02 \cdot 10^{-3}\ \mathrm{m^3/C} \pm 6{,}46 \cdot 10^{-5}\ \mathrm{m^3/C}$; P = 99% ☐
b) $A_H = 5{,}02 \cdot 10^{-3}\ \mathrm{m^3/C} \pm 5{,}69 \cdot 10^{-5}\ \mathrm{m^3/C}$; P = 99% ☐
c) $A_H = 5{,}02 \cdot 10^{-3}\ \mathrm{m^3/C} \pm 5{,}59 \cdot 10^{-5}\ \mathrm{m^3/C}$; P = 99% ☐
d) $A_H = 5{,}02 \cdot 10^{-3}\ \mathrm{m^3/C} \pm 4{,}64 \cdot 10^{-5}\ \mathrm{m^3/C}$; P = 99% ☐
e) $A_H = 5{,}02 \cdot 10^{-3}\ \mathrm{m^3/C} \pm 3{,}82 \cdot 10^{-5}\ \mathrm{m^3/C}$; P = 99% ☐

(Fragetyp Einfachwahl)

3.2. Angenommen, es sei bekannt, dass die Standardabweichung der Hall-Konstanten $\sigma_{A_H} = 8{,}5 \cdot 10^{-5}\ \mathrm{m^3/C}$ betrage. Wie groß ist dann der minimal erforderliche Stichprobenumfang n, um bei einer Aussagewahrscheinlichkeit von P = 95% das Konfidenzintervall des Erwartungswertes der Hall-Konstanten auf maximal $\pm\ 4 \cdot 10^{-5}\ \mathrm{m^3/C}$ abschätzen zu können?

a) $n = 20$ ☐
b) $n = 19$ ☐
c) $n = 18$ ☐
d) $n = 14$ ☐
e) $n = 13$ ☐

(Fragetyp Einfachwahl)

Fortsetzung Aufgabe 3 auf der nächsten Seite

3.3. Gehen Sie davon aus, dass Mittelwert und Streuung obiger Stichprobe mit dem Erwartungswert und der Standardabweichung der Grundgesamtheit übereinstimmen. Etwa wie viel Prozent aller Sensorplättchen weisen dann eine Hall-Konstante von $4{,}9 \cdot 10^{-3}\ \mathrm{m^3/C} \leq A_H \leq 5{,}1 \cdot 10^{-3}\ \mathrm{m^3/C}$ auf?

a) 90,5% ☐
b) 82,9% ☐
c) 75,3% ☐
d) 24,7% ☐
e) 17,1% ☐

(Fragetyp Einfachwahl)

3.4. Angenommen, der Erwartungswert der Hall-Konstante betrage $A_H = 5 \cdot 10^{-3}\ \mathrm{m^3/C}$. Welchen (mathematisch gerundeten) Wert dürfte die Standardabweichung σ_{A_H} der Hall-Konstante dann maximal annehmen, damit 98% der Sensorplättchen <u>innerhalb</u> des Intervalls von $4{,}9 \cdot 10^{-3}\ \mathrm{m^3/C} \leq A_H \leq 5{,}1 \cdot 10^{-3}\ \mathrm{m^3/C}$ lägen?

a) $3{,}81 \cdot 10^{-5}\ \mathrm{m^3/C}$ ☐
b) $4{,}30 \cdot 10^{-5}\ \mathrm{m^3/C}$ ☐
c) $4{,}87 \cdot 10^{-5}\ \mathrm{m^3/C}$ ☐
d) $5{,}71 \cdot 10^{-5}\ \mathrm{m^3/C}$ ☐
e) $6{,}34 \cdot 10^{-5}\ \mathrm{m^3/C}$ ☐

(Fragetyp Einfachwahl)

4. Als Hersteller von Hall-Sensoren möchten Sie den korrekten Betrieb Ihrer Fertigung sicherstellen und entnehmen zu diesem Zweck regelmäßig Stichproben aus der laufenden Produktion aller Fertigungslinien. Für einen stark nachgefragten Typ von Hall-Sensoren betreiben Sie zwei voneinander unabhängige Fertigungslinien A und B. Aufgrund eines Anfangsverdachts möchten Sie anhand der entnommenen Stichproben untersuchen, ob der Erwartungswert μ_A der Empfindlichkeit der auf Linie A gefertigten Hall-Sensoren signifikant größer ist, als der Erwartungswert μ_B der Empfindlichkeit der auf Linie B gefertigten Hall-Sensoren.

4.1. Welcher statistische Test ist geeignet, die Frage zu beantworten?

a) t-Test für Erwartungswert ☐
b) t-Test für den Vergleich zweier Erwartungswerte bei <u>un</u>abhängigen Stichproben ☐
c) t-Test für den Vergleich zweier Erwartungswerte bei verbundenen Stichproben ☐
d) F-Test für den Vergleich zweier Streuungen bei <u>un</u>abhängigen Stichproben ☐
e) χ^2-Test ☐

(Fragetyp Einfachwahl)

4.2. Welche Alternativhypothese ist für den Test zu wählen?

a) einseitige Alternativhypothese mit $\mu_A < \mu_B$ ☐
b) einseitige Alternativhypothese mit $\mu_A > \mu_B$ ☐
c) zweiseitige Alternativhypothese mit $\mu_A \neq \mu_B$ ☐

(Fragetyp Einfachwahl)

5. Anhand zweier verbundener Stichproben A und B möchten Sie einen t-Test für verbundene Stichproben durchführen. Aus den Stichproben, die jeweils einen Umfang von $n = 10$ aufweisen, haben Sie für die paarweisen Differenzen d_i einen Mittelwert von $\overline{d} = 0{,}7$ kg und eine Streuung von $S_d = 0{,}3$ kg errechnet.

5.1. Die Testgröße t_0 beträgt in diesem Fall gerundet:

a) 10,43 ☐
b) 7,38 ☐
c) 7,0 ☐
d) 1,36 ☐
e) 0,74 ☐

(Fragetyp Einfachwahl)

5.2. Der für die Bestimmung des kritischen Wertes benötigte Freiheitsgrad s beträgt bei diesem Test:

a) 20 ☐
b) 19 ☐
c) 18 ☐
d) 10 ☐
e) 9 ☐

(Fragetyp Einfachwahl)

6. Sie möchten mittels eines t-Tests für den Erwartungswert anhand einer Stichprobe die Maßhaltigkeit einer Charge von Passstiften überprüfen. Der Nenndurchmesser der Passstifte beträgt $D_{nenn} = 4$ mm. Der Stichprobenumfang beträgt $n = 20$. Ihre Nullhypothese lautet, dass der Durchmesser der Passstifte mit dem Nennwert übereinstimmt ($\mu_x = \mu_0$). Sie wählen eine zweiseitige Alternativhypothese ($\mu_x \neq \mu_0$). Sie wählen ein Signifikanzniveau von $\alpha = 0{,}01$. Die von Ihnen berechnete Testgröße beträgt $t_0 = -2{,}74$.

6.1. Geben Sie an, ob die Nullhypothese abgelehnt oder <u>nicht</u> abgelehnt werden muss!

a) Nullhypothese wird <u>nicht</u> abgelehnt ☐
b) Nullhypothese wird abgelehnt ☐

(Fragetyp Einfachwahl)

Ende von Antwort-Wahl-Verfahren, Teil A

Antwort-Wahl-Verfahren, Teil B:

7. Geben Sie an, bei welchen der folgenden Zustandsgrößen es sich um extensive Grundgrößen des SI-Systems handelt!

a) Länge ☐
b) Masse ☐
c) Volumen ☐
d) elektrischer Widerstand ☐
e) elektrische Feldstärke ☐
f) Stoffmenge ☐
g) Lichtstrom ☐
h) Temperatur ☐

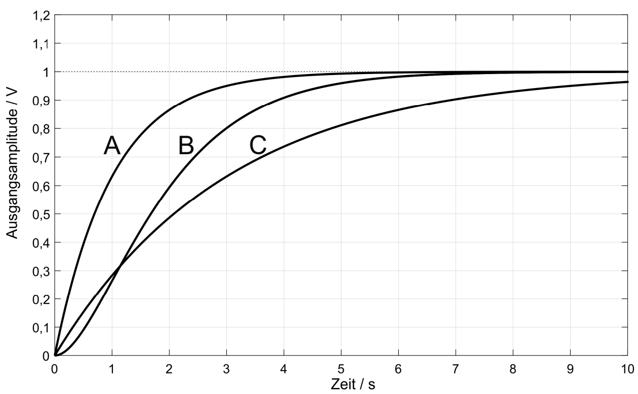
(Fragetyp Mehrfachwahl)

8. Geben Sie an, welche der folgenden Gleichungen korrekt sind!

a) $1\ \text{TW} - 100\ \text{GW} = 9{\cdot}10^{11}\ \text{W}$ ☐
b) $1\ \mu\text{V} + 1{\cdot}10^{-6}\ \text{V} = 2{\cdot}10^{-3}\ \text{mV}$ ☐
c) $1\ \text{dm}^2 + 10\ \text{cm}^2 = 1{,}1{\cdot}10^4\ \text{mm}^2$ ☐
d) $100\ \text{mV} \cdot 10\ \mu\text{A} = 1\ \mu\text{W}$ ☐
e) $100\ \text{cm}^3/\text{s} = 0{,}006\ \text{m}^3/\text{min}$ ☐

(Fragetyp Mehrfachwahl)

9. In nachfolgender Abbildung sind die Sprungantworten dreier – mit A, B und C bezeichneter – linearer Systeme dargestellt. Geben Sie für jedes der drei Systeme an, ob es sich um ein lineares System 1. Ordnung oder um ein lineares System 2. Ordnung handelt!



a) A: 1. Ordnung, B: 1. Ordnung, C: 1. Ordnung ☐
b) A: 1. Ordnung, B: 1. Ordnung, C: 2. Ordnung ☐
c) A: 1. Ordnung, B: 2. Ordnung, C: 1. Ordnung ☐
d) A: 1. Ordnung, B: 2. Ordnung, C: 2. Ordnung ☐
e) A: 2. Ordnung, B: 1. Ordnung, C: 1. Ordnung ☐
f) A: 2. Ordnung, B: 1. Ordnung, C: 2. Ordnung ☐
g) A: 2. Ordnung, B: 2. Ordnung, C: 1. Ordnung ☐
h) A: 2. Ordnung, B: 2. Ordnung, C: 2. Ordnung ☐

(Fragetyp Einfachwahl)

10. Ein lineares System 1. Ordnung mit der Zeitkonstanten T und dem Übertragungsfaktor $K = 2$ werde aus dem Beharrungszustand heraus zum Zeitpunkt $t = 0$ mit einer sprungförmigen Änderung der Eingangsspannung von $10\,\mathrm{V}$ auf $-5\,\mathrm{V}$ beaufschlagt. Welche Spannung wird nach der Zeitdauer $t = T$ am Ausgang ungefähr anliegen?

a) $-9{,}45$ V ☐
b) $-6{,}3$ V ☐
c) $0{,}55$ V ☐
d) $1{,}1$ V ☐
e) $6{,}3$ V ☐

(Fragetyp Einfachwahl)

11. Sie beobachten einen Fertigungsprozess, auf den eine große Zahl statistisch unabhängiger Einflussgrößen mit gleicher Größenordnung einwirkt. Durch welche statistische Verteilung lässt sich aller Wahrscheinlichkeit nach die Gesamtabweichung des Prozesses in guter Näherung beschreiben?

a) Gleichverteilung ☐
b) Binomialverteilung ☐
c) Hypergeometrische Verteilung ☐
d) Normalverteilung ☐
e) Poissonverteilung ☐

(Fragetyp Einfachwahl)

12. Bei dem Abtasttheorem nach Shannon handelt es sich hinsichtlich der verlustfreien Rekonstruktion der digitalisierten Daten um ein

a) hinreichendes und notwendiges Kriterium. ☐
b) hinreichendes aber <u>nicht</u> notwendiges Kriterium. ☐
c) <u>nicht</u> hinreichendes aber notwendiges Kriterium. ☐
d) <u>nicht</u> hinreichendes und <u>nicht</u> notwendiges Kriterium. ☐

(Fragetyp Einfachwahl)

13. Eine normalverteilte, dimensionslose Größe werde mit 15 Wiederholungen gemessen. Das Konfidenzintervall des Erwartungswertes wird zu $6 \leq \mu \leq 18$ bei $\mathrm{P} = 98\%$ bestimmt. Die Standardabweichung σ sei bekannt. Geben Sie an, wie viele Wiederholungsmessungen bei unveränderter Standardabweichung mindestens durchgeführt werden müssen, um das Konfidenzintervall bei unveränderter Aussagesicherheit auf $9 \leq \mu \leq 15$ zu reduzieren!

a) 22 ☐
b) 30 ☐
c) 35 ☐
d) 45 ☐
e) 60 ☐

(Fragetyp Einfachwahl)

14. Sie untersuchen anhand empirischer Daten die Studiendauer im Bachelorstudiengang Biotechnologie. Eine Auswertung der Rohdaten liefert folgende Lage- und Streuungsparameter: Der Median der Studiendauer beträgt 7,9 Semester; der Modalwert der Studiendauer beträgt 8 Semester; der arithmetische Mittelwert der Studiendauer beträgt 8,3 Semester; der Quartilsabstand der Studiendauer beträgt 1,9 Semester; das dritte Quartil der Studiendauer liegt bei 9,1 Semestern. Geben Sie an, welche der nachfolgenden Aussagen zutreffend aus diesen Daten abgeleitet werden können!

a) Die meisten Studierenden benötigen für ihr Studium 7,9 Semester. ☐

b) Die Spanne der Studiendauer beträgt 3,8 Semester. ☐

c) Die Hälfte der Studierenden benötigt bis zum Abschluss 8,3 Semester oder mehr. ☐

d) Die Hälfte der Studierenden benötigt bis zum Abschluss zwischen 7,2 und 9,1 Semester. ☐

e) Ein Viertel der Studierenden benötigt bis zum Abschluss 9,8 Semester oder mehr. ☐

(Fragetyp Mehrfachwahl)

15. Geben Sie an, welche der folgenden Aussagen über spezielle Verteilungsfunktionen zutreffend sind!

a) Die Gaußsche Normalverteilung ist symmetrisch zum Erwartungswert μ und ihre Wendepunkte liegen bei $x = \mu \pm 2\sigma$. ☐

b) Die Binomialverteilung beschreibt den wahrscheinlichen Ausgang einer Folge gleichartiger Versuche, bei der es nur zwei mögliche Ergebnisse gibt. ☐

c) Die Poissonverteilung beschreibt selten auftretende Ereignisse. Sie wurde als Grenzfall der Binomialverteilung, für eine sehr große Zahl von Versuchen $(V \to \infty)$ und eine sehr geringe Wahrscheinlichkeit des Eintretens eines Ereignisses $(p \to 0)$ hergeleitet. ☐

d) Für eine sehr große Zahl von Versuchen $(n \to \infty)$ nähert sich die Student'sche t-Verteilung der Gaußsche Normalverteilung an. ☐

e) Erwartungswert und Varianz der Poissonverteilung sind wertgleich. ☐

(Fragetyp Mehrfachwahl)

Ende von Antwort-Wahl-Verfahren, Teil B

Kurzfragen:

16. Erläutern Sie, wodurch sich *nominalskalierte* Daten und *ordinalskalierte* Daten unterscheiden! Nennen Sie für beide Datentypen je ein Beispiel!

17. Grenzen Sie die Begriffe *Messwert* und *Messergebnis* gegeneinander ab!

18. Erläutern Sie den Begriff *Repräsentativitätsfehler* und nennen Sie ein Beispiel!

19. Von der Qualitätssicherung eines Unternehmens wird mittels eines statistischen Tests überprüft, ob die produzierte Ware den geforderten Spezifikationen entspricht. Dabei wird als Nullhypothese angenommen, dass die Ware die Spezifikation erfüllt. Auf Anweisung des Managements wird die Irrtumswahrscheinlichkeit α für diesen Test von $\alpha = 0{,}05$ auf $\alpha = 0{,}01$ reduziert. In der Folge häufen sich die Beschwerden von Abnehmern, die bemängeln, dass der Anteil des Ausschusses an der gelieferten Ware zugenommen habe. Welche Erklärung haben Sie für dieses Phänomen?

Ende der Kurzfragen

Leerseite

Elementare statistische Maßzahlen

Arithmetisches Mittel: $\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n}$

Empirische Varianz: $S^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}$

Streuung: $S = +\sqrt{S^2}$

Konfidenzintervall

Die Messgröße X sei normalverteilt, σ sei bekannt:

$$\left[\bar{x} - \frac{k \cdot \sigma}{\sqrt{n}}, \bar{x} + \frac{k \cdot \sigma}{\sqrt{n}}\right]$$

Die Messgröße X sei normalverteilt, σ sei unbekannt.

$$\left[\bar{x} - \frac{S}{\sqrt{n}} t_{n-1;1-\frac{\alpha}{2}}, \bar{x} + \frac{S}{\sqrt{n}} t_{n-1;1-\frac{\alpha}{2}}\right]$$

Lineare Regression

Wenn durch eine Anzahl von Wertepaaren (x_i, y_i) nach der Methode der kleinsten quadratischen Abweichung eine Gerade gelegt wird, geht diese stets durch den Schwerpunkt $(\bar{x}, \bar{y})$ der Punkte:

$$(y - \bar{y}) = b(x - \bar{x})$$

(geschätzter) Regressionskoeffizient b (Steigung der Geraden)

$$b = \frac{\sum_{i=1}^{n}(x_i - \bar{x})\,(y_i - \bar{y})}{\sum_{i=1}^{n}(x_i - \bar{x})^2} = \frac{\sum_{i=1}^{n} x_i y_i - n\overline{xy}}{\sum_{i=1}^{n} x_i^2 - n\bar{x}^2}$$

Ein Schätzwert für σ^2 ist die Restvarianz $\hat{\sigma}^2$

$$\hat{\sigma}^2 = \frac{1}{n-2}\sum_{j=1}^{n}\left(y_j - \bar{y} + b(\bar{x} - x_j)\right)^2 = \frac{n-1}{n-2} \cdot S_y^2\left(1 - r_{xy}^2\right)$$

Bestimmung der Vertrauensgrenze für diese Schätzung des Steigungsmaßes:

1. Festlegen der geforderten statistischen Sicherheit P (z.B. 95%)
2. Berechnen der Streuung S_x aus den Messwerten $x_1, \dots, x_n$
3. Der Vertrauensbereich für den Regressionskoeffizienten b zur statistischen Sicherheit $P = 1-\alpha$ beträgt:
$$\left[b - \frac{\hat{\sigma} t_{n-2;1-\frac{\alpha}{2}}}{\sqrt{n} S_x}, b + \frac{\hat{\sigma} t_{n-2;1-\frac{\alpha}{2}}}{\sqrt{n} S_x}\right]$$
4. Der Erwartungswert β für den Regressionskoeffizienten b liegt mit der statistischen Sicherheit P in diesem Intervall
5. Durch die berechnete Gerade wird einem beliebig gewählten x-Wert x^* der y-Wert
$$y^* = \bar{y} + b(x^* - \bar{x})$$
zugeordnet. Der Vertrauensbereich für y^* zur statistischen Sicherheit $P = 1-\alpha$ beträgt:
$$\left[y^* - \frac{\hat{\sigma} t_{n-2;1-\frac{\alpha}{2}}}{\sqrt{n}}\sqrt{1 + \frac{(x^* - \bar{x})^2}{S_x^2}}, y^* + \frac{\hat{\sigma} t_{n-2;1-\frac{\alpha}{2}}}{\sqrt{n}}\sqrt{1 + \frac{(x^* - \bar{x})^2}{S_x^2}}\right]$$

Abweichungsfortpflanzung

f sei $f(x_1, \dots, x_n)$. Das Konfidenzintervall für f mit statistischer Sicherheit $P = 1-\alpha$:

$$\left[f(\bar{x}_1, \dots, \bar{x}_n) - c_f, f(\bar{x}_1, \dots, \bar{x}_n) + c_f\right]$$

für den Fall zufälliger, normalverteilter Abweichungen mit:

$$c_f = \sqrt{\sum_{i=1}^{n}\left(\left.\frac{\partial f}{\partial X_i}\right|_{\bar{x}_1,\dots,\bar{x}_n} c_{x_i}\right)^2}, c_{x_i} = \frac{S_{x_i}}{\sqrt{n_{x_i}}} t_{n_{x_i}-1;1-\frac{\alpha}{2}}$$

t-Test

t-Test für Erwartungswert

Die Testgröße:

$$t_0 = \frac{\bar{x} - \mu_0}{S/\sqrt{n}} \qquad (df = n-1)$$

Test der Nullhypothese bei vorgewähltem Signifikanzniveau α:

1. H_0: $\mu_x = \mu_0$ gegen H_1: $\mu_x < \mu_0$ (einseitige Hypothese)
Ist
$$t_0 < -t_{n-1;1-\alpha} \,,$$
wird H_0 auf dem Signifikanzniveau α abgelehnt.
2. H_0: $\mu_x = \mu_0$ gegen H_1: $\mu_x > \mu_0$ (einseitige Hypothese)
Ist
$$t_0 > t_{n-1;1-\alpha} \,,$$
wird H_0 auf dem Signifikanzniveau α abgelehnt.
3. H_0: $\mu_x = \mu_0$ gegen H_1: $\mu_x \neq \mu_0$ (zweiseitige Hypothese)
Ist
$$|t_0| > t_{n-1;1-\frac{\alpha}{2}} \,,$$
wird H_0 auf dem Signifikanzniveau α abgelehnt.

t-Test für Vergleich zweier Erwartungswerte

Die Testgröße (einfachere Form, wenn $n_x = n_y = n$):

$$t_0 = \sqrt{n}\frac{\bar{x} - \bar{y}}{\sqrt{S_x^2 + S_y^2}} \qquad (\mathrm{df} = 2\mathrm{n} - 2)$$

Test der Nullhypothese bei vorgewähltem Signifikanzniveau α:

1. H_0: $\mu_x = \mu_y$ gegen H_1: $\mu_x < \mu_y$ (einseitige Hypothese)
   Ist
   $$t_0 < -t_{n_x+n_y-2;1-\alpha} \ ,$$
   wird H_0 auf dem Signifikanzniveau α abgelehnt.
2. H_0: $\mu_x = \mu_y$ gegen H_1: $\mu_x > \mu_y$ (einseitige Hypothese)
   Ist
   $$t_0 > t_{n_x+n_y-2;1-\alpha} \ ,$$
   wird H_0 auf dem Signifikanzniveau α abgelehnt.
3. H_0: $\mu_x = \mu_y$ gegen H_1: $\mu_x \neq \mu_y$ (zweiseitige Hypothese)
   Ist
   $$|t_0| > t_{n_x+n_y-2;1-\frac{\alpha}{2}} \ ,$$
   wird H_0 auf dem Signifikanzniveau α abgelehnt.

t-Test für verbundene Stichproben

Die Testgröße:

$$t_0 = \frac{\bar{d}}{S_d/\sqrt{n}} \qquad (df = n - 1)$$

mit:

$$d_i = x_i - y_i$$

$$\bar{d} = \frac{\sum_{i=1}^{n} d_i}{n}$$

$$S_d = \sqrt{\frac{\sum_{i=1}^{n}\left(d_i - \bar{d}\right)^2}{n-1}}$$

Test der Nullhypothese bei vorgewähltem Signifikanzniveau α:

1. H_0: $\mu_d = 0$ gegen H_1: $\mu_d < 0$ (einseitige Hypothese)
   Ist
   $$t_0 < -t_{n-1,1-\alpha} \ ,$$
   wird H_0 auf dem Signifikanzniveau α abgelehnt.
2. H_0: $\mu_d = 0$ gegen H_1: $\mu_d > 0$ (einseitige Hypothese)
   Ist
   $$t_0 > t_{n-1,1-\alpha} \ ,$$
   wird H_0 auf dem Signifikanzniveau α abgelehnt.
3. H_0: $\mu_d = 0$ gegen H_1: $\mu_d \neq 0$ (zweiseitige Hypothese)
   Ist
   $$|t_0| > t_{n-1,1-\frac{\alpha}{2}} \ ,$$
   wird H_0 auf dem Signifikanzniveau α abgelehnt.

Der χ^2-Test für Verteilungsfunktionen

X sei eine Zufallsgröße mit unbekannter Verteilungsdichtefunktion. Aufgrund von Messdaten oder Vorabinformationen wird vermutet, dass X durch die Verteilungsdichtefunktion $h(x)$ beschrieben wird. Um dies zu prüfen, kann ein χ^2-Test durchgeführt werden.

Nullhypothese H_0: X wird durch die Verteilungsdichtefunktion $h(x)$ beschrieben.

Es wird eine Stichprobe von n Messwerten $x_1, \dots, x_n$ aufgenommen.

Der Test erfolgt, indem zu dieser Messreihe ein empirisches Histogramm erstellt wird. Aus der Verteilungsdichtefunktion $h(x)$ wird ein theoretisches Histogramm berechnet.

Als Testgröße wird eine normierte Differenz zwischen beiden Histogrammen berechnet. Wenn die Hypothese zutrifft, müsste diese Differenz hinreichend klein sein.

Vorgehensweise:

1. Aufteilen des Wertebereichs in r nicht überlappende Klassen T_i, so dass jede Klasse wenigstens 5 Werte der Stichprobe $x_1, \dots, x_n$ enthält. Die Intervalle können auch ungleich breit sein.
2. Bestimmen der Anzahl B_i von Messwerten in der Klasse T_i
3. Falls die Verteilungsdichtefunktion $h(x)$ Parameter enthält (z.B. μ und σ bei der Normalverteilung), so werden diese Parameter erforderlichenfalls aus den Messdaten $x_1, \dots, x_n$ abgeschätzt.
4. Berechnen der Wahrscheinlichkeit p_i, mit der bei Annahme der hypothetischen Verteilungsdichte $h(x)$ unter Annahme der unter 3. geschätzten Parameter ein Messwert im Intervall T_i zu erwarten ist.
5. Berechnen der Produkte $E_i = np_i$, die die theoretischen Besetzungszahlen der Klasse T_i bei Annahme der Verteilungsdichte $h(x)$ darstellen.
6. Prüfen, ob für alle Klassen gilt: $E_i \geq 5$. Klassen mit $E_i < 5$ werden mit benachbarten Klassen zusammengelegt. Nach diesem Schritt liegen r^* Klassen vor mit $r^* \leq r$.
7. Berechnen der Testgröße:
   $$\chi_0^2 = \sum_{i=1}^{r^*} \frac{(B_i - E_i)^2}{E_i}$$
8. Bestimmung der Zahl der Freiheitsgrade:
   - r^* ist die Zahl der auswertbaren Klassen (Besetzungszahl ≥ 5)
   - s ist die Zahl der aus der Stichprobe abgeschätzten Parameter der Verteilungsdichtefunktion
   - Die Zahl der Freiheitsgrade ist $df = r^* - s - 1$
9. Festlegen der Irrtumswahrscheinlichkeit α

H_0 ist abzulehnen mit Signifikanzniveau α, wenn:

$$\chi_0^2 > \chi_{r^*-s-1,1-\alpha}^2$$

p-Quantile $t_{s,p}$ der Student'schen t-Verteilung mit s Freiheitsgraden

s \ p	0,9	0,95	0,975	0,98	0,99	0,995
1	3,078	6,314	12,706	15,895	31,821	63,657
2	1,886	2,920	4,303	4,849	6,965	9,925
3	1,638	2,353	3,182	3,482	4,541	5,841
4	1,533	2,132	2,776	2,999	3,747	4,604
5	1,476	2,015	2,571	2,757	3,365	4,032
6	1,440	1,943	2,447	2,612	3,143	3,707
7	1,415	1,895	2,365	2,517	2,998	3,499
8	1,397	1,860	2,306	2,449	2,896	3,355
9	1,383	1,833	2,262	2,398	2,821	3,250
10	1,372	1,812	2,228	2,359	2,764	3,169
11	1,363	1,796	2,201	2,328	2,718	3,106
12	1,356	1,782	2,179	2,303	2,681	3,055
13	1,350	1,771	2,160	2,282	2,650	3,012
14	1,345	1,761	2,145	2,264	2,624	2,977
15	1,341	1,753	2,131	2,249	2,602	2,947
16	1,337	1,746	2,120	2,235	2,583	2,921
17	1,333	1,740	2,110	2,224	2,567	2,898
18	1,330	1,734	2,101	2,214	2,552	2,878
19	1,328	1,729	2,093	2,205	2,539	2,861
20	1,325	1,725	2,086	2,197	2,528	2,845
21	1,323	1,721	2,080	2,189	2,518	2,831
22	1,321	1,717	2,074	2,183	2,508	2,819
23	1,319	1,714	2,069	2,177	2,500	2,807
24	1,318	1,711	2,064	2,172	2,492	2,797
25	1,316	1,708	2,060	2,167	2,485	2,787
26	1,315	1,706	2,056	2,162	2,479	2,779
27	1,314	1,703	2,052	2,158	2,473	2,771
28	1,313	1,701	2,048	2,154	2,467	2,763
29	1,311	1,699	2,045	2,150	2,462	2,756
30	1,310	1,697	2,042	2,147	2,457	2,750
40	1,303	1,684	2,021	2,123	2,423	2,704
50	1,299	1,676	2,009	2,109	2,403	2,678
60	1,296	1,671	2,000	2,099	2,390	2,660
70	1,294	1,667	1,994	2,093	2,381	2,648
80	1,292	1,664	1,990	2,088	2,374	2,639
90	1,291	1,662	1,987	2,084	2,368	2,632
100	1,290	1,660	1,984	2,081	2,364	2,626
200	1,286	1,653	1,972	2,067	2,345	2,601
∞	1,282	1,645	1,960	2,054	2,326	2,576

p-Quantile $\chi^2_{s,p}$ der χ^2-Verteilung mit *s* Freiheitsgraden

s \ p	0,90	0,95	0,975	0,99	0,995
1	2,71	3,84	5,02	6,63	7,88
2	4,61	5,99	7,38	9,21	10,6
3	6,25	7,81	9,35	11,3	12,8
4	7,78	9,49	11,1	13,3	14,9
5	9,24	11,1	12,8	15,1	16,8
6	10,6	12,6	14,5	16,8	18,6
7	12,0	14,1	16,0	18,5	20,3
8	13,4	15,5	17,5	20,1	22,0
9	14,7	16,9	19,0	21,7	23,6
10	16,0	18,3	20,5	23,2	25,2
11	17,3	19,7	21,9	24,7	26,8
12	18,6	21,0	23,3	26,2	28,3
13	19,8	22,4	24,7	27,7	29,8
14	21,2	23,7	26,1	29,1	31,3
15	22,3	25,0	27,5	30,6	32,8
16	23,5	26,3	28,9	32,0	34,3
17	24,8	27,6	30,2	33,4	35,7
18	26,0	28,9	31,5	34,8	37,2
19	27,2	30,1	32,9	36,2	38,6
20	28,4	31,4	34,2	37,6	40,0
21	29,6	32,7	35,5	38,9	41,4
22	30,8	33,9	36,8	40,3	42,8
23	32,0	35,2	38,1	41,6	44,2
24	33,2	36,4	39,4	43,0	45,6
25	34,4	37,7	40,6	44,3	46,9
26	35,6	38,9	41,9	45,6	48,3
27	36,7	40,1	43,2	47,0	49,6
28	37,9	41,3	44,5	48,3	51,0
29	39,1	42,6	45,7	49,6	52,3
30	40,3	43,8	47,0	50,9	53,7
40	51,8	55,8	59,3	63,7	66,8
50	63,2	67,5	71,4	76,2	79,5
60	74,4	79,1	83,3	88,4	92,0
70	85,5	90,5	95,0	100,4	104,2
80	96,6	101,9	106,6	112,3	116,3
90	107,6	113,1	118,1	124,1	128,3
100	118,5	124,3	129,6	135,8	140,2

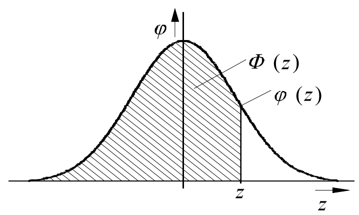
Tabelle 1

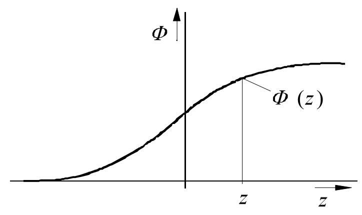
Summenfunktion der standardisierten Normalverteilung

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-t^2/2} dt \ ; \ \Phi(z) = 1 - \Phi(-z)$$

Ablesebeispiel: $\Phi(0{,}76) = 0{,}776373$

z-Transformation: $z = \dfrac{x-\mu}{\sigma}$





z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09	z
0,0	0,500000	0,503989	0,507978	0,511966	0,515953	0,519939	0,523922	0,527903	0,531881	0,535856	0,0
0,1	0,539828	0,543795	0,547758	0,551717	0,555670	0,559618	0,563559	0,567495	0,571424	0,575345	0,1
0,2	0,579260	0,583166	0,587064	0,590954	0,594835	0,598706	0,602568	0,606420	0,610261	0,614092	0,2
0,3	0,617911	0,621720	0,625516	0,629300	0,633072	0,636831	0,640576	0,644309	0,648027	0,651732	0,3
0,4	0,655422	0,659097	0,662757	0,666402	0,670031	0,673645	0,677242	0,680822	0,684386	0,687933	0,4
0,5	0,691462	0,694974	0,698468	0,701944	0,705401	0,708840	0,712260	0,715661	0,719043	0,722405	0,5
0,6	0,725747	0,729069	0,732371	0,735653	0,738914	0,742154	0,745373	0,748571	0,751748	0,754903	0,6
0,7	0,758036	0,761148	0,764238	0,767305	0,770350	0,773373	0,776373	0,779350	0,782305	0,785236	0,7
0,8	0,788145	0,791030	0,793892	0,796731	0,799546	0,802337	0,805105	0,807850	0,810570	0,813267	0,8
0,9	0,815940	0,818589	0,821214	0,823814	0,826391	0,828944	0,831472	0,833977	0,836457	0,838913	0,9
1,0	0,841345	0,843752	0,846136	0,848495	0,850830	0,853141	0,855428	0,857690	0,859929	0,862143	1,0
1,1	0,864334	0,866500	0,868643	0,870762	0,872857	0,874928	0,876976	0,879000	0,881000	0,882977	1,1
1,2	0,884930	0,886861	0,888768	0,890651	0,892512	0,894350	0,896165	0,897958	0,899727	0,901475	1,2
1,3	0,903200	0,904902	0,906582	0,908241	0,909877	0,911492	0,913085	0,914657	0,916207	0,917736	1,3
1,4	0,919243	0,920730	0,922196	0,923641	0,925066	0,926471	0,927855	0,929219	0,930563	0,931888	1,4
1,5	0,933193	0,934478	0,935745	0,936992	0,938220	0,939429	0,940620	0,941792	0,942947	0,944083	1,5
1,6	0,945201	0,946301	0,947384	0,948449	0,949497	0,950529	0,951543	0,952540	0,953521	0,954486	1,6
1,7	0,955435	0,956367	0,957284	0,958185	0,959070	0,959941	0,960796	0,961636	0,962462	0,963273	1,7
1,8	0,964070	0,964852	0,965620	0,966375	0,967116	0,967843	0,968557	0,969258	0,969946	0,970621	1,8
1,9	0,971283	0,971933	0,972571	0,973197	0,973810	0,974412	0,975002	0,975581	0,976148	0,976705	1,9
2,0	0,977250	0,977784	0,978308	0,978822	0,979325	0,979818	0,980301	0,980774	0,981237	0,981691	2,0
2,1	0,982136	0,982571	0,982997	0,983414	0,983823	0,984222	0,984614	0,984997	0,985371	0,985738	2,1
2,2	0,986097	0,986447	0,986791	0,987126	0,987455	0,987776	0,988089	0,988396	0,988696	0,988989	2,2
2,3	0,989276	0,989556	0,989830	0,990097	0,990358	0,990613	0,990863	0,991106	0,991344	0,991576	2,3
2,4	0,991802	0,992024	0,992240	0,992451	0,992656	0,992857	0,993053	0,993244	0,993431	0,993613	2,4
2,5	0,993790	0,993963	0,994132	0,994297	0,994457	0,994614	0,994766	0,994915	0,995060	0,995201	2,5
2,6	0,995339	0,995473	0,995604	0,995731	0,995855	0,995975	0,996093	0,996207	0,996319	0,996427	2,6
2,7	0,996533	0,996636	0,996736	0,996833	0,996928	0,997020	0,997110	0,997197	0,997282	0,997365	2,7
2,8	0,997445	0,997523	0,997599	0,997673	0,997744	0,997814	0,997882	0,997948	0,998012	0,998074	2,8
2,9	0,998134	0,998193	0,998250	0,998305	0,998359	0,998411	0,998462	0,998511	0,998559	0,998605	2,9

z	3,0	3,5	4,0	4,5	5,0	6,0	7,0	8,0	9,0	10,0	z
Φ(z)	$1-1{,}350\cdot10^{-3}$	$1-2{,}326\cdot10^{-4}$	$1-3{,}167\cdot10^{-5}$	$1-3{,}398\cdot10^{-6}$	$1-2{,}867\cdot10^{-7}$	$1-9{,}866\cdot10^{-10}$	$1-1{,}280\cdot10^{-12}$	$1-6{,}221\cdot10^{-16}$	$1-1{,}129\cdot10^{-19}$	$1-7{,}620\cdot10^{-24}$	Φ(z)

Φ(z)	50%	60%	70%	80%	90%	95%	97,5%	99%	99,5%	99,75%	99,9%	99,95%	Φ(z)
z	0	0,253	0,524	0,842	1,282	1,645	1,960	2,326	2,576	2,807	3,090	3,291	z