Test ◇ M2–07/08–03 ◇

Wichtig: ♡ Bitte nur die **Vorderseite** eines Blattes beschreiben.
♣ Resultate sind gut sichtbar zu unterstreichen.
♠ Nur gut leserliche, sauber gegliederte Lösungen mit sofort auffindbaren Resultaten können korrigiert werden. (Ersichtlicher Lösungsweg!)
◇ Die einzelnen Aufgaben sind durch einen Strich zu trennen.
♡ **Alle Teilaufgaben geben gleich viele Punkte.**
♣ Falls eine Frage als Fallenfrage (unsinnige Frage) erkannt wird, lautet die Antwort „Frage nicht beantwortbar".

Bootstrap–Techniken:

Probl. 1 Nenne zwei Methoden, mit welchen man exakte, nicht verteilungsfreie Vertrauensintervalle für unbekannte Parameter von Messgrössen bestimmen kann.

Probl. 2 Schildere kurz die notwendigsten, wesentlichen Eigenschaften eines Zufallsgenerators.

Probl. 3 Bei der Produktion von Plastic–Kugeln für ein Kinderspielzeug wird zwecks Qualitätssicherung eine Vorselektion gemacht. Im Selektionsprozess fallen die Kugeln zuerst durch die Maschen eines Lochsiebes mit den Lochdurchmessern von $9.001 \pm 0.001\,mm$. Anschliessend werden Kugeln ausgesondert, welche durch ein Sieb mit Lochdurchmessern von $8.799 \pm 0.001\,mm$ fallen. Damit ist ein Intervall I für die Kugeldurchmesser d_k gegeben. Die Kugeln werden schliesslich in 10-er Packungen ausgeliefert. Zur Kontrolle des Versandtgewichts und des Materialverbrauchs soll nun ein 90%–Vertrauensintervall für die Grösse $\bar{d}$ aus einem Intervall I bestimmt werden.
Schildere mit Hilfe des 5–Punkte–Schemas das Vorgehen bei der Bestimmung des Vertrauensintervalls (stichwortartig).

Probl. 4 Erkläge die Voraussetzungen, die Behauptungen und die Konsequnsequenzen für die Anwendung des Percentil–Lemmas.

Probl. 5 Bootstrap–Percentil–Vertrauensintervalle:

(a) Erkläre die Bootstrap–Methode (Plug–In–Methode) bei einem gegebenen Datensatz $M = \{x_1, x_2, \ldots, x_n\}$ für einen Parameter X. Was wird mit dieser Methode erreicht, hergestellt oder berechnet?

(b) Wie ist der Zusammenhang zwischen Plug–In–Methode und Bootstrap–Percentil–Vertrauensintervall?

(c) Was ist zum Stichprobenumfang bei der Anwendung der Plug–In–Methode zu bemerken betreffend Güte?

(d) Was ist zu den Stichprobenwerten bei der Anwendung der Plug–In–Methode zu bemerken betreffend Güte?

Probl. 6 Erkläre die Begriffe „Sensitivität", „Spezifität"und „Rate". (Erläutere die Begriffe durch Beispiele.)

Probl. 7 Erkläre den Begriff „Korrelation“. (Erläutere den Begriff durch Beispiele.)

Probl. 8 Zwei vergleichbare Datensätze sollen betreffend ihrer Streuung untersucht werden.

(a) Wie kann man sich rasch graphisch ein Bild von der Lage machen, ohne mit Verteilungsfunktionen operieren zu müssen?

(b) Schildere das Vorgehen zum Vergleich der Streuungen mit Hilfe eines „Quotienten“ und eines 95%–Percentil–Vertrauensintervalls.

Probl. 9 Was lässt sich zur Genauigkeit von Bootstrap–Percentil–Vertrauensintervallen sagen?

Probl. 10 Was lässt sich zum Aufwand bei Bootstrap–Percentil–Vertrauensintervallen bezüglich Anzahl Kopien sagen?

Probl. 11 (a) Was ist der Standardfehler einer Schätz–Statistik $\hat{\Theta}$?

(b) Was ist der Standardfehler des arithmetischen Mittelwerts einer Schätz–Statistik $\hat{\Theta}$?

(c) Nenne und erkläre eine Möglichkeit, um den Standardfehler einer relativen Häufigkeit $\neq 0$ bei Stichprobenumfang n direkt rechnerisch abzuschätzen?

Probl. 12 Wie lässt sich der Standardfehler einer Statistik empirisch mittels Bootstrapping schätzen? (Beschreibe das Verfahren!)

Probl. 13 Für eine Zufallsgrösse X liegen 410 Messwerte vor, welche ungefähr durch folgendes Modell beschrieben werden können: $x_k = 24.56 + 0.01 \cdot k + 0.02 \cdot k^2,\ \ k \in \{1, 2, 3, \ldots, 410\}$.

(a) Was ist das Besondere an dieser Art Verteilung?

(b) Jemand möchte X mit Hilfe einer Gaussverteilung modellieren. Gibt es dafür oder dagegen ein schlagendes Argument?

(c) Jemand möchte $\bar{X}$ mit Hilfe einer Gaussverteilung modellieren. Gilt das Argument aus der Antwort zur letzten Frage auch hier, wenn man die Erfahrung berücksichtigt?

Probl. 14 Gegeben ist eine Normalverteilung mit $\mu = 16.89$ und $\sigma = 0.93$. Berechne numerisch ein 99.5 %–Vertrauensintervall (Taschenrechner).

Probl. 15 Welcher mathematische Satz rechtfertigt unter günstigen Umständen gegebenenfalls die Approximation einer Verteilungen durch eine Normalverteilung?

Probl. 16 Der Erwartungswert μ_X und der Standardfehler $\widehat{SE}[\bar{X}]$ einer Messgrösse $\bar{X}$ sollen durch $\hat{\mu}_X \pm \widehat{SE}[\hat{\mu}_X] \approx 605.4 \pm 45.8$ geschätz werden. Weiter besteht zwischen X und einer von X abhängigen Messgrösse Y der Zusammenhang $Y = f(X) = X + \sin(X^2)$.

(a) Wie lautet die Formel für die 1. Approximation von Y und was ist ihr numerischer Wert?

(b) Wie lautet die Formel für die 2. Approximation von Y und was ist ihr numerischer Wert?

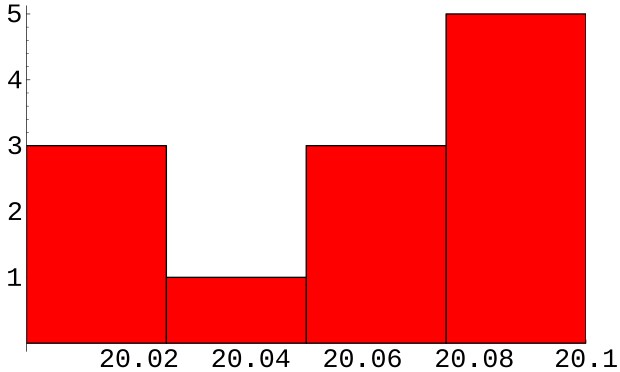
(c) Wie ist allgemein der Zusammenhang zwischen der Varianz einer Messgrösse X und ihrer transformierten Grösse Y?

(d) Berechne die Approximation der Varianz von Y bei der oben angegebenen Transformation numerisch.

Probl. 17 Gegeben ist die Datenmenge $M = \{x_1, x_2, \ldots, x_n\}$ zur Messung eines Parameters X:

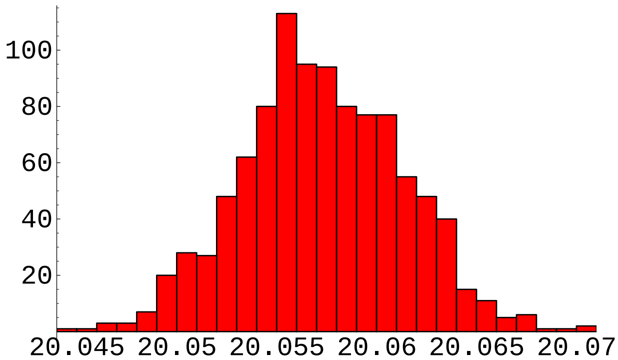
$$20.061, 20.090, 20.015, 20.031, 20.091, 20.051, 20.021, 20.081, 20.055, 20.019, 20.09, 20.080$$

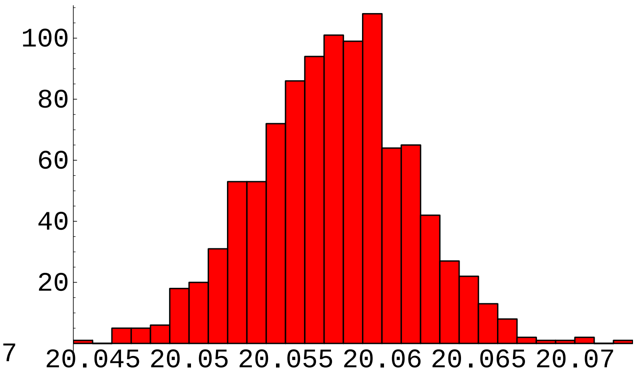
Das Histogramm dazu sieht wie folgt aus:



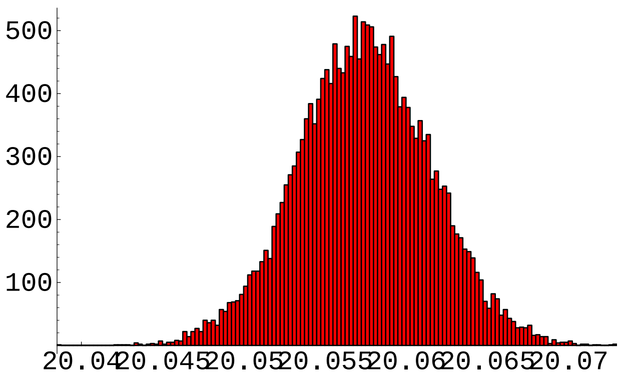
(a) Berechne den Mittelwert.

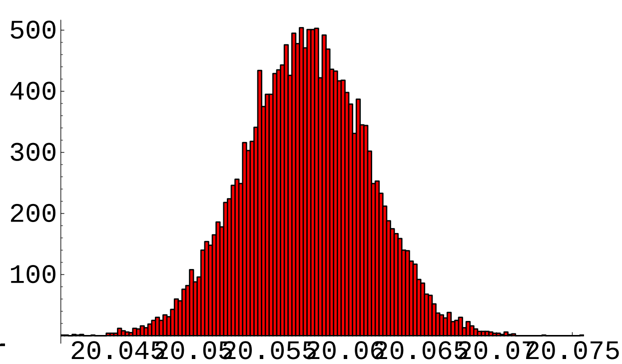
(b) Mit Hilfe eines Programms ziehen wir $n = 1000$ Bootstrap–Kopien und berechnen für jede Kopie den Mittelwert. Die gewonnenen Mittelwerte werden in eine Tabelle $TabMean$ geschrieben. Das Histogramm dazu sieht man in der Abbildung links. Erzeugen wir nochmals die Tabelle $TabMean$ mit $n = 1000$ Bootstrap–Kopien, so erhalten wir das Histogramm rechts:



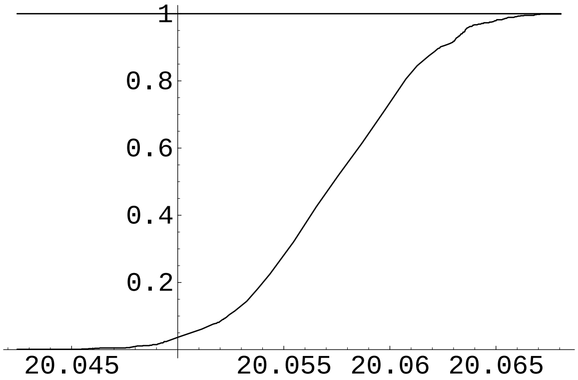


Nochmaliges Erzeugen der Tabelle $TabMean$ mit $n = 20'000$ Bootstrap–Kopien ergibt das Histogramm links. Und dann nochmals die Tabelle $TabMean$ mit $n = 20'000$ Bootstrap–Kopien erzeugen führt zum Histogramm rechts:





Um mit $n = 1000$ Kopien eine Verteilungsfunktion zu gewinnen, arbeitet der Computer zehn Minuten. Langes Warten also! Es ergibt sich das folgende Bild:



Kann man diesen nun gewonnenen Graphen als genügend genau bezeichnen, um daraus empirisch gewonnene Vertrauensintervalle für den Mittelwert ableiten zu dürfen? (Antwort begründen!)

(c) Wir nehmen hier aus pragmatischen Gründen an, dass die oben dargestellte empirisch gewonnene Verteilungsfunktion für den Mittelwert genügend genau ist. Leite daraus ein 60 %–Vertrauensintervall her, so gut es mit der Ablesegenauigkeit geht.

(d) Leite aus der eben verwendeten Graphik ein 95 %–Vertrauensintervall her, so gut es mit der Ablesegenauigkeit geht.

Viel Glück!