Vorbesprechung: 24/25. April 2013

Aufgabe 1

Die Auswertung eines Integrals

$$I(f) = \int_0^1 f(x)dx$$

kann sehr oft nicht analytisch erfolgen. Der gebräuchlichste Ansatz in diesem Fall besteht darin, das Integral numerisch zu berechnen. Dazu existieren verschiedene Computerprogramme. Eine andere geläufige Methode, um ein solches Integral zu berechnen, ist die sogenannte **Monte Carlo Methode**. Man generiert dabei uniform verteilte Zufallsvariablen auf dem Intervall $[0, 1]$, d.h., $X_1, X_2, ..., X_n$ und berechnet

$$\hat{I}(f) = \frac{1}{n}\sum_{i=1}^{n} f(X_i)\ .$$

Aufgrund des *Gesetzes der grossen Zahlen* ist für grosse n

$$\frac{1}{n}\sum_{i=1}^{n} f(X_i) \approx \mathrm{E}[f(X)]\ .$$

Somit ist

$$E[f(X)] = \int_0^1 f(x)dx = I(f)\ .$$

Dieses einfache Schema kann beliebig angepasst werden, zum Beispiel an unterschiedliche Integrationsgrenzen.
Berechnen Sie folgendes Integral

$$I(f) = \frac{1}{\sqrt{2\pi}}\int_0^1 e^{-x^2/2}dx\ .$$

Berechnen Sie das Integral, indem Sie 1000 uniform über das Intervall $[0, 1]$ verteilte Punkte, $X_1, ..., X_{1000}$ generieren. Berechnen Sie den genauen numerischen Wert des Integrals mit der R-Funktion `pnorm`.

Aufgabe 2

Aus der uniformen Verteilung $X \sim \mathrm{Uniform}([0, 10])$ soll eine Stichprobe vom Umfang n gezogen werden.

(a) Es sei $n = 60$. Bestimmen Sie ein symmetrisches Intervall $I = [\mu - e, \mu + e]$ um den Erwartungswert μ so, dass sich das arithmetische Mittel der Stichprobe mit der Wahrscheinlichkeit von 95% in I befindet. Ein solches Intervall heisst **Prognoseintervall**.

Hinweis: Standardisieren Sie das arithmetische Mittel $\overline{X}_n$ und benützen Sie den Zentralen Grenzwertsatz.

(**b**) Umgekehrt: Wie gross muss n gewählt werden, damit $e = 0.2$ wird?

(**c**) Überprüfen Sie (a) experimentell, d.h. mit `R`: ziehen Sie viele Stichproben (z.B. 200) und zählen Sie, wie viele ausserhalb von I liegen.
R-Hinweise:

```
> n<-60  # Anzahl Stichproben
> sim<-matrix(runif(n*200,min=0,max=10),ncol=n)
>        # X_1,...,X_n simulieren und in einer n-spaltigen Matrix
>        # (mit 200 Zeilen) anordnen
> sim.mean<- apply(sim,1,"mean")   #In jeder Matrixzeile Mittelwert berechnen
> plot(sim.mean)
```

Zeichnen Sie mit `abline(h=...)` die Intervallgrenzen des Prognoseintervalls in der obigen Graphik ein.

Aufgabe 3

Das Gastroberatungsunternehmen Lecker und Co. kreiert eine neue Speisekarte für ein Schnellrestaurant. Lecker und Co. nimmt auf Grund langjähriger Erfahrung an, dass etwa (unabhängig von der Anzahl der Kunden) 80% der Kunden des Schnellrestaurants die neue Speisekarte bevorzugen werden.

(**a**) Am Einführungstag speisen 356 Kunden im Restaurant. Wie gross ist unter obiger Annahme der Erwartungswert für die Anzahl der Kunden, welche die neue Speisekarte bevorzugen?

Lecker und Co. führt bei den 356 Kunden eine kurze Befragung durch. 261 Kunden geben dabei an, dass sie die neue Karte bevorzugen. Der Rest findet die alte Karte mindestens genauso gut wie die neue Karte.

(**b**) Wie gross ist die Wahrscheinlichkeit unter der Annahme von Lecker und Co., dass keiner der ersten vier befragten Kunde die neue Karte bevorzugt, der fünfte Kunde jedoch die neue Karte bevorzugt? Wie gross ist die Wahrscheinlichkeit, dass drei der ersten vier befragten Kunden die neue Karte bevorzugen?

(**c**) Wie gross ist unter der Annahme von Lecker und Co. die Wahrscheinlichkeit, dass höchstens 261 Kunden die neue Karte bevorzugen? Benutzen Sie die Normalapproximation.

(**d**) Lecker und Co. hat nun starke Zweifel, dass wirklich 80% der Kunden die neue Karte bevorzugen. Lecker und Co. will deshalb die Annahme mit Hilfe der Befragung kritisch überprüfen. Führen Sie den entsprechenden (zweiseitigen) Test durch. Verwenden Sie die Normalapproximation.

Aufgabe 4

Ein Weinhändler behauptet, dass die von ihm gefüllten Weinflaschen mindestens 70 Zentiliter enthalten. Ein skeptischer Konsument vermutet aber, dass der Weinhändler zu wenig Wein abfüllt und will diese Behauptung überprüfen. Deshalb kauft er 12 Weinflaschen und misst ihren Inhalt. Er findet:

$$71,\ 69,\ 67,\ 68,\ 73,\ 72,\ 71,\ 71,\ 68,\ 72,\ 69,\ 72 \ \text{(in Zentiliter)}.$$

(**a**) Nehmen Sie zunächst an, dass die Standardabweichung der Abfüllung im voraus bekannt ist. Sie beträgt $\sigma = 1.5$ Zentiliter. Da die Standardabweichung der Messungen bekannt ist, können wir einen z-Test durchführen. Führen Sie den (einseitigen; in welche Richtung?) Test auf dem 5%- Signifikanzniveau durch und formulieren Sie in einem Satz die Schlussfolgerung für den kritischen Konsumenten.

(**b**) Tatsächlich ist die Standardabweichung der Abfüllungen aber nicht bekannt. Man muss sie also aus den gemachten Stichproben schätzen:

$$\hat{\sigma}_x^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 \approx 1.96^2$$

Da nun die Standardabweichung geschätzt wurde und nicht mehr exakt bekannt ist, kann der z- Test nicht mehr durchgeführt werden. Verwenden Sie nun den t-Test auf dem 5%-Signifikanzniveau. Was ändert sich an obigem Test? Wie lautet das Ergebnis?

Aufgabe 5

Im National Bureau of Standards (USA) wurden regelmässig Wägungen des 10-Gramm-Standardgewichtstücks durchgeführt. Bei 9 Wägungen erhielt man als durchschnittliche Differenz -403 Mikrogramm vom 10 Gramm-Sollgewicht und eine Standardabweichung von 3.127 Mikrogramm für eine einzelne Wägung.

(**a**) Geben Sie das exakte, zweiseitige 95%-Vertrauensintervall für die wahre Differenz an, unter der Annahme, dass die Messfehler normalverteilt sind.

(**b**) Könnte die wahre Differenz $-400.0\mu g$ betragen? Entscheiden Sie aufgrund des Resultats in Aufgabe (a). (Kurze Begründung)