Dr. Christof Luchsinger

Übungsblatt 4 zur Vorlesung

"Statistische Methoden"

<u>Testtheorie: θ_0 vs θ_1</u>

Herausgabe des Übungsblattes: Woche 12, Abgabe der Lösungen: Woche 13 (bis Freitag, 1615 Uhr), Besprechung: Woche 14

Must

Aufgabe 15 [Warum Quotient und nicht Differenz der Dichten?]

Warum ist das Verhältnis der Dichten (aus $\mathcal{H}_0$ und $\mathcal{H}_1$) wichtig und nicht zum Beispiel die Differenz?

Dazu folgende 2 Hypothesen: In $\mathcal{H}_0$ haben wir die Dichtefunktion auf dem Intervall $[0, 1]$ folgendermassen konzentriert:

$$f_0(x) = \begin{cases} 8 & x \in [0, 0.05] \\ \frac{1}{9} & x \in (0.05, 0.95] \\ 10 & x \in (0.95, 1]. \end{cases}$$

In $\mathcal{H}_1$ haben wir die Dichtefunktion auf dem Intervall $[0, 1]$ folgendermassen konzentriert:

$$f_1(x) = \begin{cases} 10 & x \in [0, 0.05] \\ 0.5 & x \in (0.05, 0.95] \\ 1 & x \in (0.95, 1]. \end{cases}$$

(die Dichten können also offenbar wild verschieden sein). Wenn $\mathcal{H}_0$ richtig ist, dürfen wir in 5 % der Fälle eine Fehlentscheidung machen (Risiko 1. Art). Wie wird man sich sinnvollerweise verhalten, wenn nur eine Realisation x_1 bekannt ist (mit Satz 4.1)? Wie ist das Risiko 2. Art mit der Methode aus Satz 4.1?

Berechnen Sie in den 3 Bereichen auch die Differenzen und die Verhältnisse der beiden Dichten aus den beiden Verteilungen. Wie ist das Risiko 2. Art, wenn man auf die Differenz der Dichten schaut statt auf das Verhältnis (bei gleichem Risiko 1. Art!).

Sie werden in obigen Rechnungen eine gewisse Freiheit haben, wo Sie den Ablehnungsbereich genau wählen - aber nur eine *gewisse* Freiheit!

Aufgabe 16 [Klare Fälle und Neyman-Pearson]

Gegeben sei eine Stichprobe vom Umfang 10 aus einer Normalverteilung mit Varianz 1. Wir wissen nicht, ob der Mittelwert 0 ($\mathcal{H}_0$-Hypothese) oder 100 ($\mathcal{H}_1$-Hypothese) ist. Wie sieht ein Test mit dem Lemma von Neyman-Person aus ($\alpha = 0.1$)? Ist es sinnvoll, hier einfach stur das Lemma von Neyman-Pearson so einzusetzen?

Standard

Aufgabe 17 [Umkehrung der Fragestellung] [3+5 Punkte]

Sei $X_1, \ldots, X_n$ eine iid Folge von Be(p)-Zufallsgrössen ($P[X = 1] = p = 1 - P[X = 0]$). Eine ForscherIn möchte jetzt einen Test durchführen. Der Test sieht folgendermassen aus: Die Nullhypothese $\mathcal{H}_0 : p = 0.45$ wird genau dann abgelehnt, wenn

$$\sum_{i=1}^{n} X_i \geq n/2.$$

a) Berechnen Sie im Fall $n = 2$ die Grösse des Tests ("das α").

b) Berechnen Sie im Fall $n = 100$ die Grösse des Tests ("das α"). Benutzen Sie den CLT als approximatives Verfahren.

Aufgabe 18 [feineres Testen dank grösserem Stichprobenumfang] [3 Punkte]

Sei $x_1, \ldots, x_n$ eine Stichprobe aus einer $\mathcal{N}(\text{PN}, 1)$-Verteilung. Dabei bezeichnet PN die **P**ersonal-**N**umber jedeR StudentIn. Wir testen jetzt $\mathcal{H}_0$: Mittelwert ist (PN−0.1) gegen $\mathcal{H}_1$: Mittelwert ist PN (wir wissen, dass PN der richtige Wert ist!). Nehmen Sie $\alpha = 0.05$.

a) $n = 36$
b) $n = 100$
c) $n = 256$
d) $n = 400$
e) $n = 10'000$

Berechnen Sie zuerst in allen 5 Situationen den Ablehnungsbereich und generieren Sie danach in einer geeigneten Rechenumgebung in allen 5 Fällen eine solche Stichprobe. Wie werden Sie in diesen 5 Situationen entscheiden (wenn Sie kurz vergessen, dass *Sie* wissen, dass PN der richtige Mittelwert ist)?

Honours

Aufgabe 19 [mit Hilfe von R/S-PLUS; Bsp wo nicht MLQ gilt] [1+2 Punkte]

Die Cauchy-Zufallsgrösse (vgl. 1.4.2.5) ist ein praktisches Gegenbeispiel für viele Untersuchungen ($E[|X|] = \infty$ und vieles mehr). Die Dichtefunktion ist

$$f(x) = \frac{d}{\pi(d^2 + (x - m)^2)};$$

dabei ist m der Median und d ein Skalenparameter. Wir setzen hier $d = 1$ und untersuchen mit einer Einerstichprobe ($n = 1$), ob $m = 0$ ($\mathcal{H}_0$) oder $m = 1$ ($\mathcal{H}_1$). Die minimal suffiziente Statistik ist $x := x_1$. Wir wollen (*und können*!) Satz 4.1 anwenden. Schwierig wird (wegen fehlendem MLQ) die Berechnung des Ablehnungsbereichs.

a) Untersuchen Sie als Vorbereitung auf b), wie sich der Likelihood-Quotient verhält (wo fallend, steigend, wieder fallend; keine genauen Berechnungen, sondern grobe Abschätzung reicht).

b) Berechnen Sie in R/S-PLUS durch pröbeln die Grenzen, wo Sie *mit Satz 4.1* die Nullhypothese ablehnen / Alternativhypothese annehmen sollten. Nehmen Sie $\alpha = 0.1$ und suchen Sie Werte, sodass die Genauigkeit 5 Promille beträgt (Risiko erster Art im Intervall $[9.5, 10.5]$). Es wird klar verlangt, dass Satz 4.1 benutzt wird und also das Risiko 2. Art minimiert wird. Wir suchen nicht irgendein Intervall oder Bereich, wo wir $\mathcal{H}_0$ ablehnen, sondern den Bereich, damit das Risiko 2. Art minimal ist. Tipp: a<-seq(0,3,0.01) und b<-dcauchy(a,1)/dcauchy(a,0); Vorsicht: Indexe um 1 verschoben ([1] $\equiv$ 0.00 und *nicht* 0.01)!

Übungsblatt 4 zur Vorlesung "Statistische Methoden"

Olivier Warin

5. April 2011

Aufgabe 15 [Warum Quotient und nicht Differenz der Dichten?]
Sei $x_1 \in [0,1]$ eine Stichprobe aus einer stetigen Zufallsgrösse X. Nun wollen wir die folgenden beiden Hypothesen gegeneinander auf dem Niveau $\alpha = 5\%$ testen:

$$\mathcal{H}_0 : \text{ Die Dichte von } X \text{ hat auf dem Intervall } [0,1] \text{ die Form } f_0(x) = \begin{cases} 8, & x \in [0, 0.05] \\ \frac{1}{9}, & x \in (0.05, 0.95] \\ 10, & x \in (0.95, 1]. \end{cases}$$

$$\mathcal{H}_1 : \text{ Die Dichte von } X \text{ hat auf dem Intervall } [0,1] \text{ die Form } f_1(x) = \begin{cases} 10, & x \in [0, 0.05] \\ 0.5, & x \in (0.05, 0.95] \\ 1, & x \in (0.95, 1]. \end{cases}$$

- Zunächst testen wir mit der Methode aus Satz 4.1. Dazu bestimmen wir zuerst einmal den Quotienten der beiden Dichten:

$$\frac{f_1(x)}{f_0(x)} = \begin{cases} 1.25, & x \in [0, 0.05] \\ 4.5, & x \in (0.05, 0.95] \\ 0.1, & x \in (0.95, 1]. \end{cases}$$

Wenn wir dem Satz 4.1 strikt folgen wollten, so müssten wir ein $K \in \mathbb{R}$ finden, so dass

$$\alpha = P_0 \left[\frac{f_1(X)}{f_0(X)} > K \right].$$

Nun kann man leicht einsehen, dass dies hier nicht möglich ist. Wir wählen den Ablehnungsbereich daher wie folgt:

$$(0.05, 0.5),$$

denn es gilt

$$\alpha \; = \; 0.05 = \int_{0.05}^{0.5} \frac{1}{9} dx \; = \; \int_{0.05}^{0.5} f_0(x) dx \; = \; P_0[X \in (0.05, 0.5)] \; = \; P_0 \left[\frac{f_1(X)}{f_0(X)} > 4, X < 0.5 \right].$$

Also haben wir *fast* die von Satz 4.1 gewünschte Form gefunden.
Das Risiko zweiter Art β lautet nun also wie folgt:

$$\beta = P_1[X \notin (0.05, 0.5)] \; = \; \int_0^{0.05} f_1(x) dx + \int_{0.5}^1 f_1(x) dx \; = \; 0.775.$$

- Nun testen wir indem wir die Differenz der Dichten anschauen:

$$f_1(x) - f_0(x) = \begin{cases} 2, & x \in [0, 0.05] \\ 7/18, & x \in (0.05, 0.95) \\ -9, & x \in (0.95, 1]. \end{cases}$$

Ähnlich wie oben wählen wir nun den Ablehnungsbereich wie folgt:

$$(0, 0.00625),$$

denn es gilt

$$\alpha = 0.05 = \int_0^{0.00625} 8dx = \int_0^{0.00625} f_0(x)dx = P_0[X \in (0, 0.00625)]$$
$$= P_0[f_1(X) - f_0(X) > 1, X < 0.00625].$$

Wir haben also fast die analoge Form (mit Differenz statt Quotient) wie in Satz 4.1.

Mit diesem Ablehnungsbereich lautet das Risiko zweiter Art β wie folgt:

$$\beta = P_1[X > 0.00625] = \int_{0.00625}^{1} f_1(x)dx = 0.9375.$$

Wenn wir mit der Differenz statt mit dem Quotient der Dichten arbeiten erhalten wir also ein deutlich grösseres Risiko erster Art.

Aufgabe 16
Gegeben sei eine Stichprobe $\boldsymbol{x} = (x_1, \ldots, x_{10})$ vom Umfang 10 aus einer Normalverteilung mit Varianz 1 und Erwartungswert μ. Wir testen jetzt auf dem Niveau $\alpha = 0.1$ mit dem Lemma von Neyman-Person die folgenden zwei Hypothesen gegeneinander:

$$\begin{aligned} \mathcal{H}_0 : &\quad \mu = 0 \\ \mathcal{H}_1 : &\quad \mu = 100 \end{aligned}$$

Die entsprechende gemeinsame Dichtefunktion lautet wie folgt:

$$f(\boldsymbol{x}) \stackrel{\text{II}}{=} \prod_{i=1}^{10} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x_i-\mu)^2} = \frac{1}{32\pi^5} \exp\left(-\frac{1}{2} \sum_{i=1}^{10} (x_i - \mu)^2 \right).$$

Somit hat der Likelihood-Quotient die folgende Form:

$$\frac{f_1(\boldsymbol{x})}{f_0(\boldsymbol{x})} = \exp\left(\frac{1}{2} \sum_{i=1}^{10} (200x_i - 100) \right) = \exp(1000\overline{\boldsymbol{x}} - 500).$$

Nach dem Lemma von Neyman-Person (Satz 4.1) brauchen wir nun ein $K \in \mathbb{R}$ mit

$$0.1 = \alpha = P_0[\exp(1000\overline{\boldsymbol{X}} - 500) > K] = P_0\left[\overline{\boldsymbol{X}} > \underbrace{\frac{\log K}{1000} + \frac{1}{2}}_{=:K'} \right].$$

Natürlich reicht es, wenn wir das K' bestimmen. Dies können wir schnell tun, da $\overline{\boldsymbol{X}}$ unter $\mathcal{H}_0$ eine $\mathcal{N}(0, 1/10)$-Verteilung hat: tun

$$K' \underset{\text{R}}{\doteq} \texttt{qnorm(0.1,0,sqrt(1/10),lower.tail=FALSE)} \underset{\text{R}}{\doteq} 0.4052622.$$

Also werden wir $\mathcal{H}_0$ ablehnen, sobald $\overline{\boldsymbol{x}} > K' \doteq 0.4052622$.

Bemerkung: In dieser Situation ist es natürlich nicht so sinnvoll stur das Lemma von Neyman-Pearson so einzusetzen. Denn man kann sofort erkennen, ob $\mathcal{H}_0$ oder $\mathcal{H}_1$ anzunehmen ist. Wenn man nach dem Lemma von Neyman-Pearson vorgeht, geht man einfach ein Risiko erster Art ein ohne das Risiko 2. Art entsprechend zu reduzieren.

Aufgabe 17 [Umkehrung der Fragestellung]
Sei $X_1, \ldots, X_n$ eine iid Folge von Be(p)-Zufallsgrössen ($P[X_1 = 1] = p = 1 - P[X_1 = 0]$). Eine ForscherIn möchte jetzt einen Test durchführen. Der Test sieht folgendermassen aus: Die Nullhypothese $\mathcal{H}_0 : p = 0.45$ wird genau dann abgelehnt, wenn

$$\sum_{i=1}^{n} X_i \geqslant \frac{n}{2}.$$

a) Nehmen wir an, dass $n = 2$ gilt. Nun folgt für die Grösse des Tests α:

$$\alpha = P_0\left[\sum_{i=1}^{n} X_i \geqslant \frac{n}{2}\right] = P_0[X_1 + X_2 \geqslant 1] = 1 - P_0[X_1 + X_2 < 1] = 1 - P_0[X_1 = 0, X_2 = 0]$$
$$\overset{\text{II}}{=} 1 - P_0[X_1 = 0]P_0[X_2 = 0] = 1 - (1 - 0.45)(1 - 0.45) = 0.6975.$$

b) Hier nehmen wir an, dass $n = 100$ gilt. Jetzt approximieren wir die Grösse des Tests α mit Hilfe des CLTs:

$$\alpha = P_0\left[\sum_{i=1}^{n} X_i \geqslant \frac{n}{2}\right] = P[\text{Bin}(100, 0.45) \geqslant 50]$$
$$= P\left[\frac{\text{Bin}(100, 0.45) - 100 \cdot 0.45}{\sqrt{100 \cdot 0.45 \cdot (1 - 0.45)}} \geqslant \frac{50 - 100 \cdot 0.45}{\sqrt{100 \cdot 0.45 \cdot (1 - 0.45)}}\right] \underset{\text{CLT}}{\doteq} P[\mathcal{N}(0,1) \geqslant 5/\sqrt{24.75}]$$
$$\underset{\text{R}}{\doteq} \texttt{pnorm(5/sqrt(24.75), lower.tail=FALSE)} \underset{\text{R}}{\doteq} 0.1574393.$$

Bemerkung: Mit Hilfe von R können wir hier auch auf den CLT verzichten und erhalten so das folgende, etwas genauere Ergebnis:

$$\alpha = P_0\left[\sum_{i=1}^{n} X_i \geqslant \frac{n}{2}\right] = P[\text{Bin}(100, 0.45) \geqslant 50] \underset{\text{R}}{\doteq} \texttt{pbinom(50-1,100,0.45,lower.tail=FALSE)}$$
$$\underset{\text{R}}{\doteq} 0.1827282$$

Aufgabe 18 [feineres Testen dank grösserem Stichprobenumfang]
Sei $x_1, \dots, x_n$ eine Stichprobe aus einer $\mathcal{N}(\text{PN}, 1)$-Verteilung. Dabei bezeichnet PN die **P**ersonal **N**umber jedeR StudentIn. In den Musterlösungen werden wir PN = 2 verwenden.

Wir testen jetzt auf dem Niveau $\alpha = 0.05$ $\mathcal{H}_0$: Mittelwert ist $(\text{PN} - 0.1)$ gegen $\mathcal{H}_1$: Mittelwert ist PN.

Zuerst bestimmen wir dazu den Ablehnungsbereich A_n (in Abhängigkeit von n). Nach dem Lemma von Neyman-Person hat dieser Ablehnungsbereich die folgende Form:

$$A_n = \left\{(x_1, \cdots, x_n) \in \mathbb{R}^n \mid \frac{f_1(x_1, \cdots, x_n)}{f_0(x_1, \cdots, x_n)} > K_n\right\},$$

wobei K_n eine bestimmte reelle Zahl ist.

Wie im Skript auf Seite 69 können wir dieses A_n aber viel einfacher schreiben (beachte dazu, dass $PN - 0.1 < PN$ gilt):

$$A_n = \{\boldsymbol{x} \in \mathbb{R}^n \,|\, \overline{\boldsymbol{x}} > K_n'\}, \tag{$*$}$$

wobei K_n' eine bestimmte reelle Zahl bezeichnet.

Da wir auf dem Niveau $\alpha = 0.05$ testen wollen, muss K_n' die folgende Gleichung erfüllen:

$$0.05 = \alpha = P_0[\overline{\boldsymbol{X}} > K_n'] = P[\mathcal{N}(\text{PN} - 0.1, 1/n) > K_n'] \underset{\text{formation}}{\overset{\text{Z-Trans-}}{=}} P\left[\mathcal{N}(0,1) > \sqrt{n}(K_n' - \text{PN} + 0.1)\right].$$

Es folgt $\sqrt{n}(K_n' - \text{PN} + 0.1) \underset{\text{R}}{\doteq}$ `qnorm(0.05,lower.tail=FALSE)` und damit

$$K_n' \underset{\text{R}}{\doteq} \frac{1}{\sqrt{n}}\texttt{qnorm(0.05,lower.tail=FALSE)} + \text{PN} - 0.1.$$

Also wird K_n' für $n \to \infty$ immer kleiner und geht gegen $\text{PN} - 0.1$. D.h. wenn n grösser wird, nehmen wir tendenziell eher $\mathcal{H}_1$, was ja richtig ist.

Aufgrund von $(*)$ reicht es nun für die Ablehnungsbereiche jeweils das zugehörige K_n' anzugeben:

a) $n = 36$, also $K_n' \underset{\text{R}}{\doteq}$ `qnorm(0.05,lower.tail=FALSE)/sqrt(36)+1.9` $\underset{\text{R}}{\doteq} 2.174142$.

Eine Simulation in R lieferte den folgenden Wert für $\overline{\boldsymbol{x}}$:

```
> mean(rnorm(36,2,1))
[1] 1.976065
```

Da $1.976065 < K'_n$ nehmen wir hier (fälschlicherweise) $\mathcal{H}_0$ an.

b) $n = 100$, also $K'_n \underset{\text{R}}{\doteq}$ `qnorm(0.05,lower.tail=FALSE)/sqrt(100)+1.9` $\underset{\text{R}}{\doteq} 2.064485$.

Eine Simulation in R lieferte den folgenden Wert für $\overline{\boldsymbol{x}}$:

```
> mean(rnorm(100,2,1))
[1] 1.989742
```

Da $1.989742 < K'_n$ nehmen wir hier ebenfalls (fälschlicherweise) $\mathcal{H}_0$ an.

c) $n = 256$, also $K'_n \underset{\text{R}}{\doteq}$ `qnorm(0.05,lower.tail=FALSE)/sqrt(256)+1.9` $\underset{\text{R}}{\doteq} 2.002803$.

Eine Simulation in R lieferte den folgenden Wert für $\overline{\boldsymbol{x}}$:

```
> mean(rnorm(256,2,1))
[1] 2.081600
```

Da $2.081600 > K'_n$ lehnen wir hier $\mathcal{H}_0$ ab.

d) $n = 400$, also $K'_n \underset{\text{R}}{\doteq}$ `qnorm(0.05,lower.tail=FALSE)/sqrt(400)+1.9` $\underset{\text{R}}{\doteq} 1.982243$.

Eine Simulation in R lieferte den folgenden Wert für $\overline{\boldsymbol{x}}$:

```
> mean(rnorm(400,2,1))
[1] 2.056102
```

Da $2.056102 > K'_n$ lehnen wir hier $\mathcal{H}_0$ ab.

e) $n = 10'000$, also $K'_n \underset{\text{R}}{\doteq}$ `qnorm(0.05,lower.tail=FALSE)/sqrt(10000)+1.9` $\underset{\text{R}}{\doteq} 1.916449$.

Eine Simulation in R lieferte den folgenden Wert für $\overline{\boldsymbol{x}}$:

```
> mean(rnorm(10000,2,1))
[1] 2.009805
```

Da $2.009805 > K'_n$ lehnen wir hier $\mathcal{H}_0$ ab.

Aufgabe 19 [mit Hilfe von R/S-PLUS; Bsp wo nicht MLQ gilt]
In dieser Aufgabe geht es um eine Cauchy-Zufallsgrösse, also um eine stetige Zufallsgrösse mit der folgenden Dichtefunktion:

$$f(x) = \frac{d}{\pi(d^2 + (x-m)^2)};$$

dabei ist m der Median und d ein Skalenparameter. Hier setzen wir $d = 1$ und untersuchen mit einer Einerstichprobe x_1 ($n = 1$), ob $m = 0$ ($\mathcal{H}_0$) oder $m = 1$ ($\mathcal{H}_1$) gilt. Eine minimal suffiziente Statistik ist $x = x_1$.

a) Als Vorbereitung betrachten wir den entsprechenden Likelihood-Quotienten:

$$\begin{aligned} g(x) &:= \frac{f_1(x)}{f_0(x)} = \frac{1+x^2}{1+(x-1)^2} \\ \Rightarrow\ g'(x) &= \frac{2x(1+(x-1)^2) - 2(x-1)(1+x^2)}{(x^2+1)^2}. \end{aligned}$$

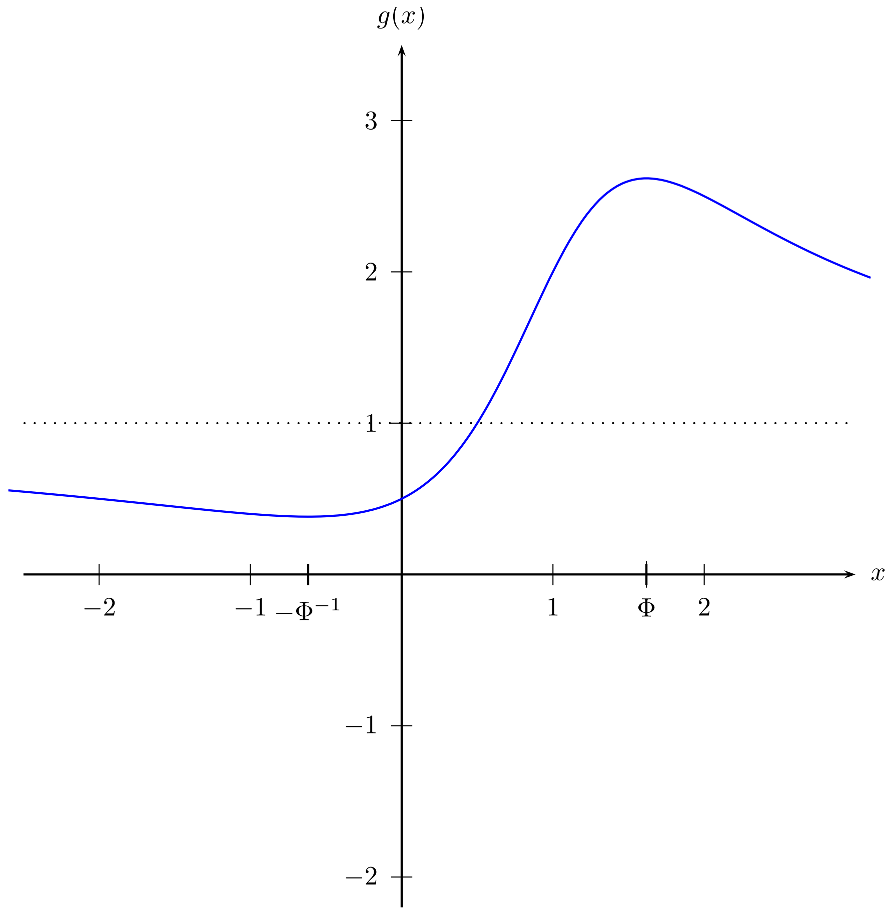
Damit folgt:

$$\begin{aligned} g'(x) \lesseqqgtr 0 \quad &\Leftrightarrow \quad 2x(1+(x-1)^2) - 2(x-1)(1+x^2) \lesseqqgtr 0 \\ &\Leftrightarrow \quad x^2 - x - 1 = (x-\Phi)(x+\Phi^{-1}) \gtreqqless 0, \end{aligned}$$

wobei $\Phi = \frac{1+\sqrt{5}}{2}$ (und damit $-\Phi^{-1} = \frac{1-\sqrt{5}}{2}$) die goldene Zahl bezeichnet.

Wir schliessen, dass g auf $(-\infty, -\Phi^{-1})$ und auf (Φ, ∞) streng monoton fallend und auf $(-\Phi^{-1}, \Phi)$ streng monton wachsend ist.

Zur besseren Vorstellung ist hier eine Skizze des Graphen:



b) **Bemerkung:** Die Idee dieser Aufgabe war es den Ablehnungsbereich durch Pröbeln herauszufinden. Da sich aber pröbeln schlecht in eine Musterlösung einbauen lässt, ist hier eine algebraische Lösung der Aufgabe. Wie man die Aufgabe durch Pröbeln lösen kann, erfahren Sie in der Übungsstunde.

Nach a) können wir die Funktion g jeweils auf den Intervallen $(-\infty, -\Phi^{-1})$, $(-\Phi^{-1}, \Phi)$ und (Φ, ∞) umkehren. Wir führen nun die folgenden drei Bezeichungen ein: Die Umkehrabbildung von $g\big|_{(-\infty,-\Phi^{-1})}$ bezeichnen wir mit ψ_1, die Umkehrabbildung von $g\big|_{(-\Phi^{-1},\Phi)}$ bezeichnen wir mit ψ_2 und die Umkehrabbildung von $g\big|_{(\Phi,\infty)}$ bezeichnen wir mit ψ_3.

Laut Satz 4.1 hat der gesuchte Ablehnungsbereich die Form $\{x \in \mathbb{R} \,|\, g(x) > K\}$, wobei K eine reelle Zahl mit der folgenden Eigenschaft ist:

$$0.1 = \alpha = P_0[g(X) > K].$$

Mit Hilfe von a) bestimmen wir schnell $\min_{x\in\mathbb{R}} g(x) = g(-\Phi^{-1}) = \frac{3-\sqrt{5}}{2}$ und $\max_{x\in\mathbb{R}} g(x) = g(\Phi) = \frac{3+\sqrt{5}}{2}$. Daraus folgt, dass sicher

$$\frac{3-\sqrt{5}}{2} < K < \frac{3+\sqrt{5}}{2}$$

gelten muss. Denn sonst wäre $P_0[g(X) > K]$ 0 bzw. 1 und somit sicher nicht gleich α.

Wir vermuten aber gleich noch eine etwas stärkere Bedingung, nämlich

$$1 < K < \frac{3+\sqrt{5}}{2}.$$

Dies wird sich auch als richtig herausstellen, daher nehmen wir diese Bedingung an. Denn diese Annahme vereinfacht die Rechnung etwas. Wenn wir hier falsch liegen würden, so müssten wir auf einen Widerspruch kommen und müssten die Annahme verwerfen.

Beachten wir, dass ψ_2 monoton wachsend und ψ_3 monoton fallend ist, schliessen wir:

$$\begin{aligned}\alpha &= P_0[g(X) > K] \\ &= \underbrace{P_0[g(X) > K, X < -\Phi^{-1}]}_{=0 \text{ (da } g \text{ hier } < 1 \text{ ist)}} + P_0[g(X) > K, -\Phi^{-1} < X < \Phi] + P_0[g(X) > K, \Phi < X] \\ &= P_0[\psi_2(K) < X < \Phi] + P_0[\Phi < X < \psi_3(K)] \\ &= P_0[X < \Phi] - P_0[X < \psi_2(K)] + P_0[\psi_3(K)] - P_0[X < \Phi] \\ &= P_0[\psi_2(K) < X < \psi_3(K)].\end{aligned}$$

Also hat der gesuchte Ablehnungsbereich die Form (a,b), wobei $a = \psi_2(K)$ und $b = \psi_3(K)$. Nun müssen wir noch a und b bestimmen. Dazu beachten wir zuerst, dass gilt

$$g(a) = g(\psi_2(K)) = K = g(\psi_3(K)) = g(b)$$

und damit

$$\frac{1+a^2}{1+(1-a)^2} = \frac{1+b^2}{1+(1-b)^2}$$

woraus durch etwas umformen folgt:

$$b = \frac{a+2}{2a-1}.$$

Zu erfüllen bleibt also die Gleichung

$$\begin{aligned}\alpha = P_0[a < X < b] &= P_0\left[a < X < \frac{a+2}{2a-1}\right] = \int_a^{\frac{a+2}{2a-1}} f_0(x)dx = \frac{1}{\pi}\int_a^{\frac{a+2}{2a-1}} \frac{1}{x^2+1}dx \\ &= \frac{1}{\pi}\left(\arctan\left(\frac{a+2}{2a-1}\right) - \arctan(a)\right).\end{aligned}$$

Daraus folgt

$$\pi\alpha + \arctan(a) = \arctan\left(\frac{a+2}{2a-1}\right).$$

Mit der Bezeichnung $t = \tan(\pi\alpha)$ und der trigonometrischen Formel $\tan(x+y) = \frac{\tan x + \tan y}{1 - \tan x \tan y}$ erhalten wir damit

$$\frac{t+a}{1-ta} = \frac{a+2}{2a-1}.$$

Unter Beachtung, dass $a > -\Phi^{-1}$, führt uns dies zu

$$\begin{aligned}a &= \frac{-(2t-1)+\sqrt{5(t^2+1)}}{t+2} \underset{\text{R}}{\doteq} 1.161890289 \\ b &= \frac{a+2}{2a-1} \underset{\text{R}}{\doteq} 2.388530504.\end{aligned}$$

Der gesuchte Ablehnungsbereich lautet also wie folgt:

$$(a,b) \doteq (1.161890289, 2.388530504).$$