Übungsaufgaben zur Vorlesung

Multiple Testtheorie

Humboldt-Universität zu Berlin
Wintersemester 2013/2014

Dr. Thorsten Dickhaus
Weierstraß-Institut für Angewandte Analysis
und Stochastik (WIAS) Berlin
www.math.hu-berlin.de/~dickhaus/
Thorsten.Dickhaus@wias-berlin.de
Tel.: 030/20372-453

Aufgaben

1. **Strukturierte Hypothesensysteme**

   Konstruieren Sie je ein Hypothesensystem, in welchem

   (a) die Menge der Maximalhypothesen ungleich der Menge der Globalhypothesen ist.

   (b) die Menge der Minimalhypothesen ungleich der Menge der Elementarhypothesen ist.

2. **Kohärenz multipler Tests**

   (a) Beweisen Sie Lemma 1.17 aus der Vorlesung.

   (b) Beweisen Sie Teil (a) von Lemma 1.18 aus der Vorlesung.

3. **Erwartete Anzahl von Typ I Fehlern**

   Für einen multiplen Test φ für das multiple Testproblem $(\Omega, \mathcal{A}, \mathcal{P}, \mathcal{H})$ mit $\mathcal{H} = \{H_i, i \in I = \{1, \ldots, m\}\}$ bezeichne die Zufallsgröße

   $$V_\varphi(\vartheta) := \sum_{i \in I_0(\vartheta)} \varphi_i$$

   die Anzahl der fälschlicherweise verworfenen Nullhypothesen. Beweisen Sie die folgenden Ungleichungen:

   $$\forall \vartheta \in \Theta : \frac{\mathbb{E}_\vartheta[V_\varphi(\vartheta)]}{m_0} \leq \mathbb{P}_\vartheta\Big(\bigcup_{i \in I_0(\vartheta)} \{\varphi_i = 1\}\Big) \leq \mathbb{E}_\vartheta[V_\varphi(\vartheta)].$$

   Dabei sei wie üblich $m_0 = |I_0(\vartheta)|$ die Anzahl wahrer Nullhypothesen.

4. **Durchführung eines multiplen Tests in Handrechnung**

   In einer Untersuchung zum Fettgehalt in sogenannten „Light"-Butterprodukten wird für drei Buttersorten der Fettgehalt Y in g pro 100g festgestellt. Von jeder Sorte werden vier Proben betrachtet. Zusätzliche Effekte auf den Fettgehalt der Proben bleiben unberücksichtigt. Zur Analyse wird das lineare Modell $y_{ij} = \beta_i + \varepsilon_{ij}$, $i = 1, 2, 3$, $j = 1, \ldots, 4$ mit $\varepsilon_{ij} \sim \mathcal{N}(\mu, \sigma^2)$ iid., $\sigma^2 > 0$ unbekannt, unterstellt. Folgende Ergebnisse werden notiert:

   Fettgehalt in g je 100g Butter

Sorte 1	Sorte 2	Sorte 3
61	62	65
58	59	62
60	61	63
60	61	62

   (a) Stellen Sie das durchschnittsabgeschlossene Hypothesensystem zu der Fragestellung auf, ob zwischen den Sorten bezüglich des Fettgehalts Unterschiede bestehen. Testen Sie jede der resultierenden Nullhypothesen zum lokalen Niveau $\alpha = 0.05$.

(b) Treffen Sie eine Aussage zur Widerspruchsfreiheit der erzielten Testentscheidung.

(c) Was lässt sich über Kohärenz und Konsonanz dieser Entscheidung sagen?

5. **Programmieraufgabe**

In dem Artikel Notterman, D. A., Alon, U., Sierk, A. J. (2001). Transcriptional Gene Expression Profiles of Colorectal Adenoma, Adenocarcinoma, and Normal Tissue Examined by Oligonucleotide Arrays. Cancer Research 61, 3124-3130, finden sich publizierte Daten aus einem Krebs-Forschungsprojekt. Das Ziel der Studie war, differentiell exprimierte Gene und R(D)NA-Profile in Tumorgewebe im Vergleich mit normalem (gesundem) Gewebe zu finden.

Dazu wurde eine klinische Studie mit $n = 22$ Krebspatienten durchgeführt. Wir betrachten hier nur den Teildatensatz der 18 Patienten mit Adenokarzinom. Von diesen 18 Individuen wurden Genexpressionsdaten („Intensitäten") für 7457 verschiedene RNA-, DNA- und Genorte erhoben, und zwar jeweils einmal in einer Gewebeprobe mit Tumor und einmal in einer gesunden Gewebeprobe. Die zugehörigen (leicht aufbereiteten) Daten sind Bestandteil des `mutoss`- Zusatzpakets für `R`.

In solchen QTL- (quantitative trait loci) Analysen wird typischerweise eine Log-Normalverteilung für die Intensitätsquotienten angenommen. Nach einigen Vorverarbeitungsschritten (siehe „Materials and Methods" in dem o.a. Artikel) wurden daher zum Vergleich der beiden Gruppen gepaarte t-Tests für verbundene Stichproben auf den durch den natürlichen Logarithmus transformierten Daten vorgeschlagen.

(a) Laden Sie die Nutzdaten in `R`. Vollziehen Sie die p-Werte für die zweiseitigen t-Tests nach.

(b) Welche RNA-, DNA- bzw. Genorte zeigen nach Bonferroni-Adjustierung eine zum multiplen Niveau $\alpha = 0.05$ signifikante differentielle Expression in Tumor- im Vergleich zu gesundem Gewebe? Geben Sie die entsprechenden Indizes an.

6. **Abschlusstest**

Vergegenwärtigen Sie sich noch einmal die Problematik der Typ III Fehler aus Beispiel 1.31. Nehmen Sie zu dem in Beispiel 1.31 aufgestellten Hypothesensystem noch die Hypothesen $H_< : \{\mu_1 < \mu_2\}$ sowie $H_> : \{\mu_1 > \mu_2\}$ hinzu. Das erweiterte Hypothesensystem ist dann also gegeben als $\mathcal{H} = \{H_<, H_\leq, H_=, H_\geq, H_>\}$.

(a) Stellen Sie die Hypothesenstruktur von $\mathcal{H}$ schematisch dar. Ermitteln Sie die bestehenden Obermengenbeziehungen.

(b) Geben Sie geeignete Tests zum lokalen Niveau α für die beiden neu hinzu genommenen Hypothesen $H_<$ und $H_>$ an und bilden Sie den Abschlusstest für $\mathcal{H}$.

(c) Was kann nunmehr aus einer Realisierung der Teststatistik inferiert werden? Stellen Sie die Entscheidungsbereiche auf der reellen Achse graphisch dar.

7. **Simultane Güte**

Betrachten Sie zur Bewertung eines multiplen Tests $\varphi = (\varphi_1, \ldots, \varphi_m)$ die Wahrscheinlichkeit, dass φ insgesamt eine korrekte Entscheidung trifft. Diese ist für $\vartheta \in \Theta$ gegeben als

$$PC_\varphi(\vartheta) = \mathbb{P}_\vartheta(\varphi = \varepsilon(\vartheta)) = \mathbb{P}(\bigcap_{i \in I} \{\varphi_i = \varepsilon_i(\vartheta)\}) \text{ mit } \varepsilon_i(\vartheta) = \begin{cases} 0, & \vartheta \in H_i \\ 1, & \vartheta \in K_i \end{cases}, \ i = 1, \ldots, m.$$

Zeigen Sie: Ist φ ein Test zum multiplen Niveau α, so gilt

$$PC_\varphi(\vartheta) \leq SG_\varphi(\vartheta) \leq PC_\varphi(\vartheta) + \alpha,$$

mit der simultanen Güte $SG_\varphi(\vartheta)$ wie in Definition 1.37.

8. **p-Wert für den einseitigen Binomialtest**

   Man nehme an, es soll die Zuverlässigkeit von Transistoren geprüft werden. Der Hersteller macht die Angabe, dass ein neuer Transistor mit einer Wahrscheinlichkeit von mindestens $\pi_0 = 99.9\%$ ein Jahr lang unter standardisierten Bedingungen ausfallfrei betrieben werden kann. In einem großen Labor werden tausend solcher Transistoren unabhängig voneinander ein Jahr lang unter den vom Hersteller genannten standardisierten Bedingungen betrieben. Drei der Transistoren fallen dabei aus. Spricht dies signifikant gegen die Herstellergarantie? Zur Beantwortung dieser Frage bezeichne man die wahre Nichtausfallwahrscheinlichkeit eines Transistors des Herstellers in einem Jahr unter den standardisierten Bedingungen mit π und berechne den p-Wert für das Testproblem

   $$H_0 : \{\pi \geq \pi_0\} \quad \text{versus} \quad H_1 : \{\pi < \pi_0\}.$$

   Hinweis: Die Anzahl ausfallender Transistoren in einem Jahr unter den standardisierten Bedingungen in einer Stichprobe vom Umfang n ist unter der Annahme, dass die n Transistoren unabhängig voneinander betrieben werden, binomialverteilt mit Parametern n und $1 - \pi$.

9. **Bedingte Momente von bootstrap Größen**

   Unter den Voraussetzungen von Beispiel 2.9 sei $\vec{X} = (X_1, \ldots, X_n)$ der Vektor der iid. Originaldaten der erhobenen Zufallsstichprobe.

   Berechnen Sie die folgenden beiden bedingten Momente der zugehörigen iid. Variablen eines bootstrap Datensatzes.

   $$\text{(a)} \;\; \mathbb{E}(X_1^* | \vec{X}), \qquad \text{(b)} \;\; \mathrm{Var}(X_1^* | \vec{X}).$$

10. **Programmieraufgabe**

    Wir betrachten das statistische Modell $(\mathbb{R}, \mathcal{B}(\mathbb{R}), (\mathcal{N}(\vartheta, 1))_{\vartheta \in \mathbb{R}_{\geq 0}})$, wobei $\mathcal{B}(\mathbb{R})$ die Borel'sche σ-Algebra auf $\mathbb{R}$ bezeichnet. Zu testen sei

    $$H_0 : \{\vartheta = 0\} \quad \text{versus} \quad H_1 : \{\vartheta > 0\}. \tag{1}$$

    Dazu liege die folgende Stichprobe vom Umfang $n = 15$ vor: $1.311, 1.136, 1.81, 0.827, -0.173, 0.351, -1.949, 0.973, 0.617, -0.091, -0.155, -0.581, 0.452, 0.879, 0.17$.

    (a) Berechnen Sie probabilistisch den exakten p-Wert für das Testproblem (1) basierend auf der obigen Stichprobe. Verwenden Sie dazu die in `R` verfügbare Funktion zur Auswertung der Verteilungsfunktion der Standardnormalverteilung.

    (b) Implementieren Sie den im Vorlesungsskript (Kapitel 2) vorgestellten Bootstrap. Ermitteln Sie den bootstrap p-Wert für das Testproblem (1) basierend auf der obigen Stichprobe und wählen Sie dabei $B = 9999$. Vergleichen Sie das Ergebnis mit dem Ergebnis aus Aufgabenteil (a).

11. **Monotonie von Likelihood Ratio-basierten Testfamilien**

    Sei X ein Zufallsvektor mit Werten in $\Omega \subset \mathbb{R}^k$ und mit der Dichte f_ϑ bezüglich eines σ-finiten Maßes μ für $\vartheta \in \Theta \subset \mathbb{R}$. Ferner sei eine Testfamilie $(\mathcal{H}, \mathcal{T})$ mit $\mathcal{H} = \{H_i, i \in I^*\}$, $\mathcal{T} = \{T_i, i \in I^*\}$ und $I^* = \{0, 1, \ldots, m\}$ gegeben.

    Zeigen Sie: Falls alle Teststatistiken $T_i, i \in I^*$ auf Likelihood Ratio-Statistiken beruhen, so ist $(\mathcal{H}, \mathcal{T})$ monoton im Sinne von Definition 3.2. Dabei ist eine Likelihood Ratio-Statistik für ein Testproblem H_i versus K_i, $i \in I^*$ mit $H_i + K_i = \Theta$ gegeben als eine monoton fallende Funktion des folgenden Quotienten:

    $$\sup_{\vartheta \in H_i} f_\vartheta(x) / \sup_{\vartheta \in \Theta} f_\vartheta(x).$$

12. **Erweiterter Korrespondenzsatz**

    Beweisen Sie den erweiterten Korrespondenzsatz 3.9.

13. **GT2-Test von Hochberg und Scheffé-Test in Handrechnung**

    Von Interesse sei der Vergleich des durchschnittlichen Gewichts dreier Apfelsorten. Leider ist bei einem Sturm fast die komplette Ernte vernichtet worden und Sie finden von den Sorten „Sonnenschein“ und „Morgenröte“ jeweils nur vier brauchbare Äpfel und von der Sorte „Jonas“ fünf brauchbare Exemplare. Deren Gewichte sind in der folgenden Tabelle notiert:

    Apfelgewicht in g

Sonnenschein	Morgenröte	Jonas
97.1	99.3	102.5
99.5	100.3	100.9
98.8	101.3	101.6
98.3	99.0	101.8
		101.4

    Mögen die Zufallsvariablen $X_{ij}, i = 1, 2, 3, j = 1, \dots, n_i$ mit $n_i = 4 + \mathbf{1}\{i = 3\}$ die Apfelgewichte beschreiben und es werde Modell 3.10 für $X = (X_{ij}, i = 1, 2, 3, j = 1, \dots, n_i)$ angenommen.

    Führen Sie unter Zuhilfenahme von Tafeln für die „maximum modulus“- und die F-Verteilung alle drei paarweisen Sortenmittelwertsvergleiche in Handrechnung zum multiplen Niveau $\alpha = 0.05$ durch

    (a) mit der GT2-Methode von Hochberg.

    (b) mit dem Scheffé-Test.

    (c) Vergleichen Sie die unter (a) und (b) gewonnenen Aussagen.

14. **Dunnett- und Tukey-Test in Handrechnung**

    Im Rahmen einer landwirtschaftlichen Versuchsreihe sei man an Kohlkopferträgen bei Verwendung verschiedener Düngemittel interessiert. Dazu seien $X_{ij}, i = 1, \dots, 4, j = 1, \dots, 5$ Zufallsvariablen, die die Ernteerträge des j-ten Kohlkopfes in der zu Dünger i gehörigen Stichprobe beschreiben. Für $X = X_{ij}, i = 1, \dots, 4, j = 1, \dots, 5$ werde Modell 3.10 angenommen. Das erhobene Datenmaterial ist in der folgenden Tabelle zusammengefasst.

    Kohlkopfgewicht in g

Dünger 1	Dünger 2	Dünger 3	Dünger 4
728	974	548	450
955	748	682	405
823	937	763	529
1161	904	617	759
972	869	632	397

    (a) Nehmen Sie an, dass Düngemittel 1 ein Standardprodukt ist, während die anderen neu entwickelte Produkte darstellen. Testen Sie mit dem „many-one t-Test“ von Dunnett zum multiplen Niveau $\alpha = 0.01$ in Handrechnung, ob sich zwischen dem Standardprodukt und jedem einzelnen der drei neu entwickelten Produkte hinsichtlich des mittleren Kohlkopfertrags ein Unterschied feststellen lässt.

    (b) Führen Sie den Tukey-Test zum multiplen Niveau $\alpha = 0.01$ zum paarweisen Vergleich aller vier Mittelwerte durch. Vergleichen Sie die damit gewonnenen Aussagen für die Fragestellung unter (a) mit denen, die der „many-one t-Test“ von Dunnett zum multiplen Niveau $\alpha = 0.01$ geliefert hat.

15. **Programmieraufgabe**

Betrachten Sie das bereits in der Vorlesung andiskutierte Beispiel von Keuls (1952, Euphytica 1, 112-122). Führen Sie unter Zuhilfenahme der R-Software alle 78 paarweisen Sortenmittelwertsvergleiche zum multiplen Niveau $\alpha = 0.05$ durch

    (a) mit dem Bonferroni t-Test.
    (b) mit dem Scheffé-Test.
    (c) mit dem Tukey-Test.
    (d) Machen Sie Aussagen zum Gütevergleich der drei unter (a), (b) und (c) angewendeten Methoden anhand dieses Datensatzes.

16. **Vergleich des Abschlusstests mit Bonferroni-Holm**

Nehmen Sie Modell 3.10 mit $k = 3$ an.

    (a) Geben Sie die Testvorschrift des Bonferroni-Holm Tests für $\mathcal{H} = (H_{ij}, 1 \leq i < j \leq 3)$ unter Verwendung der t-Statistiken $T_{ij}, 1 \leq i < j \leq 3$, an.
    (b) Bilden Sie den Abschlusstest für $\mathcal{H}$, wobei der Test für die Schnitthypothese durch Bonferroni-Adjustierung gebildet werde.
    (c) Welcher der unter (a) und (b) konstruierten multiplen Tests für $\mathcal{H}$ ist (hinsichtlich Güte) besser? Geben Sie ein Beispiel an, in dem sich einer der beiden Tests (bezüglich der Anzahl abgelehnter Hypothesen) als überlegen erweist.
    (d) Betrachten Sie das Hypothesensystem $\tilde{\mathcal{H}} = (H_i, i = 1, 2, 3)$ mit $H_i : \{\mu_i = 0\}$. Was lässt sich jetzt zum Vergleich von Bonferroni-Holm Test und Abschlusstest mit Bonferroni-Adjustierung aussagen?

17. **Verbesserung von Bonferroni-Holm durch logische Restriktionen**

Sei $(\Omega, \mathcal{A}, \mathcal{P}, \mathcal{H} = (H_i, i \in I = \{1, \ldots, m\}))$ ein endliches multiples Testproblem mit $m^* \leq m$ Elementarhypothesen. Betrachten Sie den step-down Test, der analog zu Bonferroni-Holm arbeitet, aber die kritischen Werte

$$\alpha_\ell = \frac{\alpha}{t_\ell}, \ell = 1, \ldots, m^*,$$

verwendet, wobei t_ℓ die maximal mögliche Anzahl wahrer Hypothesen, gegeben, dass $\ell - 1$ Hypothesen falsch sind, bezeichnet.

    (a) Begründen Sie, warum diese Prozedur ein multipler Test zum multiplen Niveau α ist.
    (b) Zeigen Sie, dass diese Prozedur (hinsichtlich Güte) besser als Bonferroni-Holm ist.
    (c) Geben Sie unter Modell 3.10 mit $k = 4$ für $\mathcal{H} = (H_{ij}, 1 \leq i < j \leq 4)$ die Werte der t_ℓ an.

18. **LSD-Test von Fisher und Bonferroni-Holm Test in Handrechnung**

Betrachten Sie noch einmal die Problemstellung aus Aufgabe 14 (Düngemittelvergleiche).

    (a) Führen Sie alle paarweisen Mittelwertvergleiche zwischen den Düngemittelgruppen mit dem LSD-Test von Fisher für $\alpha = 0.01$ durch.
    (b) Führen Sie alle paarweisen Mittelwertvergleiche zwischen den Düngemittelgruppen mit dem Bonferroni-Holm Test für $\alpha = 0.01$ durch.
    (c) Vergleichen Sie die unter (a) und (b) gewonnenen Aussagen mit denen, die der Tukey-Test aus Aufgabe 14.(b) geliefert hat.

19. **Multiple range-Test nach Newman und Keuls in Handrechnung**

Unter Modell 3.10 mit balanciertem Design und $k = n = 5$ seien die Gruppenmittel

$$\bar{x}_{1.} = 20.7, \quad \bar{x}_{2.} = 17.0, \quad \bar{x}_{3.} = 16.1, \quad \bar{x}_{4.} = 21.1, \quad \bar{x}_{5.} = 26.5$$

beobachtet worden. Der Schätzwert für die gepoolte Stichproben-Standardabweichung sei $s = 2.683$. Führen Sie zum Vergleich der Gruppenmittel den multiple range-Test nach Newman und Keuls mit $\alpha = 0.05$ durch.

20. **Programmieraufgabe**

Betrachten Sie noch einmal das Beispiel von Notterman et al. aus Aufgabe 5.

(a) Führen Sie unter Zuhilfenahme der R-Software alle 7457 gepaarten t-Tests mit dem Bonferroni-Holm Test zum multiplen Niveau $\alpha = 0.05$ durch.

(b) Verifizieren Sie mit Hilfe der R-Software, dass ein step-up Test mit kritischen Werten $\alpha_i = \alpha/(10i)$, $i = 1, \ldots, 7457$, ein multipler Test zum multiplen Niveau $\alpha = 0.05$ für die Fragestellung unter (a) ist.

(c) Führen Sie den unter (b) vorgeschlagenen step-up Test am Computer durch und vergleichen Sie ihn mit dem Bonferroni-Holm Test hinsichtlich der Anzahl abgelehnter Hypothesen.

21. **Schätzung des Anteils wahrer Nullhypothesen**

Als Verallgemeinerung des Storey-Schätzers betrachten Sie den Schätzer $\hat{m}_0(\kappa, \lambda)$ für die Anzahl wahrer Nullhypothesen in einem endlichen Hypothesensystem $\mathcal{H} = \{H_i, i \in I = \{1, \ldots, m\}\}$, der gegeben ist durch

$$\hat{m}_0(\kappa, \lambda) = m \frac{1 - \hat{F}_m(\lambda) + \kappa/m}{1 - \lambda}.$$

Dabei sind $\kappa \geq 0$ und $\lambda \in (0, 1)$ Tuningparameter und $\hat{F}_m$ bezeichnet wie in der Vorlesung die empirische Verteilungsfunktion der marginalen p-Werte $(p_i(X), i \in I)$.

(a) Berechnen Sie unter dem Dirac-uniform Modell $\mathrm{DU}_{m_0,m}$ mit m_0 wahren und $m - m_0$ falschen Nullhypothesen den Erwartungswert von $\hat{m}_0(\kappa, \lambda)$.

(b) Nehmen Sie an, Sie wollen den Bonferroni-Test (bzw. den Šidák-Test) verbessern, indem Sie deren kritischen Wert für die marginalen p-Werte von α/m zu $\alpha/\hat{m}_0(\kappa, \lambda)$ bzw. von $1 - (1 - \alpha)^{1/m}$ zu $1 - (1 - \alpha)^{1/\hat{m}_0(\kappa,\lambda)}$ vergrößern. Wie muss κ jeweils gewählt werden, damit keine Hypothese abgelehnt werden kann, deren zugehöriger marginaler p-Wert den Wert λ übersteigt?

22. **Per Family Error Rate**

Betrachten Sie die Per Family Error Rate (PFER) aus Definition 1.33, also die (unter $\vartheta \in \Theta$) erwartete Anzahl an Typ I-Fehlern eines multiplen Tests $\varphi = (\varphi_i, i \in I = \{1, \ldots, m\})$.

(a) Berechnen Sie unter den generellen Voraussetzungen (D2) - (I2) aus Definition 5.6 $\mathrm{PFER}_\vartheta(\varphi^{(k)})$, wobei $\varphi^{(k)}$ ein Einschritttest mit dem kritischen Wert α/k für die marginalen p-Werte $(p_i(X), i \in I)$ mit Parameter $k \geq 1$ ist. Dabei sei $\vartheta \in \Theta$ so beschaffen, dass genau m_0 der m zu testenden Hypothesen wahr sind.

(b) Wie kann k unter den obigen Voraussetzungen gewählt werden, so dass $\varphi^{(k)}$ die PFER zum Niveau $\alpha \in (0, 1)$ kontrolliert?

23. **Storey-Methode in Handrechnung**

Gegeben sei ein multiples Testproblem $(\Omega, \mathcal{A}, \mathcal{P}, \mathcal{H} = \{H_i, i = 1, \ldots, 30\})$, für das die generellen Voraussetzungen (D2) - (I2) erfüllt seien. Die folgenden 30 marginalen p-Werte seien beobachtet worden.

0.4598367567,	0.0219807061,	0.2828991192,	0.6783085944,	0.7365392294,
0.9014605558,	0.0539129250,	0.3089559420,	0.7103366164,	0.1864614866,
0.8255097728,	0.9522515490,	0.3928713611,	0.3850308341,	0.7490618181,
0.0660193717,	0.1813801406,	0.0676278590,	0.0304835552,	0.1323320201,
0.0411958267,	0.0079257922,	0.3540481755,	0.0419575551,	0.0056959610,
0.0049364264,	0.0031589772,	0.1442576172,	0.5328830657,	0.6059783014.

Führen Sie mit Hilfe der Storey-Methode mit $\lambda = 0.5$ eine FDR-basierte Analyse des obigen Datensatzes zum FDR-Niveau $\alpha = 0.05$ durch.

24. **Step-up-down basierend auf der asymptotisch optimalen Ablehnkurve**

Unter den Voraussetzungen von und unter Verwendung der Daten aus Aufgabe 23 führen Sie einen step-up-down Test basierend auf der asymptotisch optimalen Ablehnkurve $f_{0.05}$ mit Tuningparameter $\lambda = \alpha = 0.05$ in Handrechnung durch.

Vergleichen Sie Ihr Ergebnis mit dem aus Aufgabe 23.

25. **Programmieraufgabe**

Betrachten Sie noch einmal das Beispiel von Notterman et al. aus Aufgabe 5. Führen Sie unter Zuhilfenahme der `R`-Software alle 7457 gepaarten t-Tests zum FDR-Niveau $\alpha = 0.05$ durch

(a) mit dem linearen step-up Test von Benjamini und Hochberg.

(b) mit der Storey-Prozedur.

(c) mit einem auf der asymptotisch optimalen Ablehnkurve f_α basierenden SUD-Test mit Parameter $\lambda = \alpha$.