

Faglig kontakt under eksamen:
Bo Lindqvist 975 89 418

EKSAMEN I FAG TMA4275 LEVETIDSANALYSE

Mandag 27. mai 2013

Tid: 09:00–13:00

Tillatte hjelpemidler:

Godkjent kalkulator.

Ett gult ark (A4 med stempel) med dine egne formler og notater.

Sensur: 17. juni 2013.

BOKMÅL

Tabeller over χ^2 -fordelingen med 1 frihetsgrad og standardnormalfordelingen er gitt til slutt i oppgavesettet.

Oppgave 1 *Behandlinger for prostatakraft*

En randomisert kontrollert klinisk studie for å sammenligne behandlinger for prostatakraft ble påbegynt i 1967 av Veterans Administration Cooperative Urological Research Group. To av behandlingene som ble brukt i studien var et placebo og 1,0 mg med dietylstilbestrol (DES), begge administrert daglig gjennom munnen. Starttidspunkt i studien er datoen da en pasient ble randomisert til en behandling, og endepunktet er pasientens død av prostatakraft. Overlevelsestider for pasienter som døde av andre årsaker, eller som ble tapt under oppfølgingsprosessen, ble ansett som sensurerte.

Dataene som brukes i denne oppgaven er fra en bestemt undergruppe av pasienter med kreft som har nådd et visst stadium. Dataene er gitt til slutt i oppgaven.

Følgende variabler er registrert for hver pasient:

Treat: *Behandlingsgruppe.* 0 = placebo; 1 = DES.

Time: *Overlevelsestid,* muligens sensurert (måneder).

Status: *Sensureringsstatus*. 0 = sensurert; 1 = død.

Age: *Alder* (år).

Shb: *Serum haemoglobin-nivå* (gm/100 ml).

Size: *Størrelse på svulst* (cm²).

Index: *Gleason-indeks* (en kombinert indeks for svulstens stadium og grad; jo lenger svulsten har kommet, jo høyere er verdien på indeksen).

- a) Bruk dataene til å beregne Kaplan-Meier estimatoren separat for pasientene som fikk placebo (Treat = 0) og pasienter som fikk DES (Treat = 1), uten å ta de andre kovariatene i betraktning. (Merk at DES-gruppen bare har én observert død).

Tegn de to kurvene i samme figur.

Hvilken foreløpig konklusjon kan trekkes fra denne figuren?

Hva er estimert medianoverlevelse og hva er estimert forventet overlevelse for de pasientene som fikk placebo? Hvorfor kan ikke medianoverlevelse beregnes for DES-pasientene? Kan den estimerte forventede overlevelse for DES-pasientene beregnes? Hvis ja, hva er estimatet?

De fire prognostiske variablene Age, Shb, Size og Index kan ha en effekt på overlevelsestidene, og man utfører derfor en Weibull-regresjon ved hjelp av disse kovariatene i tillegg til Treat-variabelen. Utskriften fra MINITAB er som følger:

Regression Table

Predictor	Coef	Standard Error	Z	P	95,0% Normal CI	
					Lower	Upper
Intercept	8,97979	3,32964	2,70	0,007	2,45383	15,5058
Treat	0,425215	0,459101	0,93	0,354	-0,474606	1,32504
Age	-0,0070739	0,0231046	-0,31	0,759	-0,0523580	0,0382103
Shb	-0,0505088	0,148834	-0,34	0,734	-0,342217	0,241200
Size	-0,0368758	0,0170023	-2,17	0,030	-0,0701996	-0,0035520
Index	-0,275063	0,119843	-2,30	0,022	-0,509952	-0,0401752
Shape	2,73540	0,955769			1,37913	5,42547

Log-Likelihood = -31,341

- b) Sett opp den komplette modellen bak denne utskriften. Bruk kovariatnavnene gitt i utskriften (se beskrivelse i begynnelsen av oppgaven). Sett også opp et uttrykk for median overlevelsestid for en pasient som funksjon av parametrene i modellen og de observerte kovariatene.

Gi en tolkning av de estimerte regresjonskoeffisientene med hensyn på betydningen av den tilsvarende kovariat på overlevelsestiden for en pasient.

Hva er estimert relativ økning i median overlevelsestid hvis en pasient mottar DES-behandling istedenfor placebo? (Beregn først kvotienten mellom estimert median overlevelsestid under DES og under placebo).

Hvilke av forklaringsvariablene har signifikant effekt? Viser resultatet en signifikant effekt av å bruke DES? (Bruk signifikansnivå 5% under undersøkelsen av kovariater).

Det ble besluttet å bruke bare Treat, Size og Index som kovariater i de neste analysene. For å undersøke betydningen av behandling (Treat), gjorde man én analyse med Treat i modellen og én uten Treat i modellen. Her er resultatene fra MINITAB:

Predictor	Coef	Standard Error	Z	P	95,0% Normal CI	
					Lower	Upper
Intercept	7,73138	1,45451	5,32	0,000	4,88059	10,5822
Treat	0,434133	0,463267	0,94	0,349	-0,473854	1,34212
Size	-0,0370421	0,0174203	-2,13	0,033	-0,0711853	-0,0028989
Index	-0,269215	0,116184	-2,32	0,020	-0,496931	-0,0414985
Shape	2,69158	0,939226			1,35826	5,33377

Log-Likelihood = -31,434

Predictor	Coef	Standard Error	Z	P	95,0% Normal CI	
					Lower	Upper
Intercept	8,21515	1,54411	5,32	0,000	5,18876	11,2415
Size	-0,0446092	0,0179154	-2,49	0,013	-0,0797227	-0,0094958
Index	-0,289745	0,124149	-2,33	0,020	-0,533072	-0,0464182
Shape	2,71880	0,956222			1,36459	5,41690

Log-Likelihood = -32,015

- c) La β_1 være koeffisienten til variabelen Treat i modellen som tilsvare den første av de to utskriftene ovenfor. For å teste om det er en effekt av DES, betrakter man følgende hypoteser:

$$H_0 : \beta_1 = 0 \text{ mot } H_1 : \beta_1 \neq 0.$$

Det er (minst) to formelle måter å teste dette på basert på utskriftene ovenfor, nemlig å bruke "Coef" eller å bruke "Log-likelihood".

Du skal beskrive begge måtene, og finne de respektive p -verdiene (for en p -verdi tatt direkte fra MINITAB-utskriften, må du gi en forklaring på hvordan den beregnes).

Row	Treat	Time	Status	Age	Shb	Size	Index
1	0	2	0	76	10,7	8	9
2	0	14	1	73	12,4	18	11
3	0	23	0	68	12,5	2	8
4	0	24	0	71	13,7	10	9
5	0	26	1	72	15,3	37	11
6	0	36	1	72	16,4	4	9
7	0	42	1	57	13,9	24	12
8	0	43	0	60	13,6	7	9
9	0	51	0	61	13,5	8	8
10	0	52	0	73	11,7	5	9
11	0	58	0	64	16,2	6	9
12	0	59	0	77	12,0	7	10
13	0	61	0	75	13,7	10	12
14	0	62	0	63	13,2	3	8
15	0	65	0	67	13,4	34	8
16	0	67	0	70	14,7	7	9
17	0	67	0	71	15,6	8	8
18	0	69	1	60	16,1	26	9

19	1	5	0	74	15,1	3	9
20	1	16	0	73	13,8	8	9
21	1	28	0	75	13,7	19	10
22	1	45	0	72	11,0	4	8
23	1	50	1	68	12,0	20	11
24	1	51	0	65	14,1	21	9
25	1	51	0	65	14,4	10	9
26	1	54	0	51	15,8	7	8
27	1	55	0	74	14,3	7	10
28	1	57	0	72	14,6	8	10
29	1	60	0	77	15,6	3	8
30	1	61	0	60	14,6	4	10
31	1	64	0	74	14,2	4	6
32	1	65	0	51	11,8	2	6
33	1	66	0	70	16,0	8	9
34	1	66	0	70	14,5	15	11
35	1	67	0	73	13,8	7	8
36	1	68	0	71	14,5	19	9
37	1	70	0	72	13,8	3	9
38	1	70	0	71	13,6	2	10

Oppgave 2 *Utskifting av ventilseter*

I en flåte av dieselmotorer antas at utskiftingen av ventilseter for en enkelt motor følger en homogen Poisson-prosess (HPP) med (konstant) intensitet λ . Intensiteten antas å ha samme verdi for alle motorer.

- a) Betrakt først én slik motor. La $N(t)$ være antall utskiftinger i tidsintervallet fra 0 til $t > 0$.

Hva er fordelingen for $N(t)$ for en gitt tid t ?

Hva er forventet antall utskiftinger i tidsintervallet fra 0 til t ?

Hva er sannsynligheten for at det er nøyaktig én utskifting av ventilsete i tidsintervallet fra 0 til t ?

Hva er fordelingen for tiden til den første utskifting for en motor? Vis hvordan du kan finne denne fordelingen fra kjennskap til fordelingen for $N(t)$ for $t > 0$.

Anta i det følgende at det er m motorer i flåten, der den j te motoren er observert i tidsintervallet $(0, \tau_j)$, med N_j utskiftinger i dette intervallet ($j = 1, \dots, m$).

- b) Utled likelihood-funksjonen (rimelighetsfunksjonen) $L(\lambda)$ under de gitte modellforutsetningene og med de gitte tellingene av ventilseteutskiftinger. Vis at denne funksjonen kan uttrykkes som

$$L(\lambda) = \lambda^{\sum_{j=1}^m N_j} e^{-\lambda \sum_{j=1}^m \tau_j},$$

hvor \ln er den naturlige logaritme.

Finn sannsynlighetsmaksimeringsestimatorens $\hat{\lambda}$ for λ .

- c) Utled en estimator for standardavviket til $\hat{\lambda}$, og et tilnærmet 95% konfidensintervall for λ .

Beregn $\hat{\lambda}$, standardavvikestimaten og konfidensintervallet når $m = 3$ og dataene er gitt som følger:

j	τ_j	N_j
1	240	3
2	300	1
3	450	2

Oppgave 3 *Den log-logistiske fordeling*

La T være levetiden for et viss type utstyr, modellert ved en log-lokasjon-skala familie, slik at

$$\ln T = \mu + \sigma W,$$

der μ og σ er parametere ($\sigma > 0$), mens W har den standard logistiske fordeling med kumulativ fordelingsfunksjon (CDF)

$$\Phi(w) = \frac{e^w}{1 + e^w} \text{ for } -\infty < w < \infty.$$

T sies nå å ha den *log-logistiske fordeling* med parametere μ og σ .

a) Vis at CDF for T er gitt ved

$$F(t) = \frac{e^{\frac{\ln t - \mu}{\sigma}}}{1 + e^{\frac{\ln t - \mu}{\sigma}}} \text{ for } t > 0. \quad (1)$$

La t_p være $100p$ -prosentilen i fordelingen til T , for $0 < p < 1$, dvs. tiden slik at $F(t_p) = p$.
Vis at

$$\ln t_p = \mu + \sigma \ln \frac{p}{1-p}.$$

Hva er medianen i fordelingen for T ? Finn også et uttrykk for interkvartilbredden for T , dvs. forskjellen mellom 75- og 25-prosentilene for T .

Anta nedenfor at det foreligger et høyresensurert datasett (Y_i, δ_i) , $i = 1, \dots, n$, for et utvalg fra det gitte utstyret. Her er Y_i en observert tid mens δ_i er sensureringsstatus; $\delta_i = 1$ hvis Y_i er en levetid; $\delta_i = 0$ hvis Y_i er en sensureringstid.

b) Forklar hvordan uttrykket for $F(t)$ i (1) kan brukes til å konstruere et sannsynlighetsplott for å sjekke om datasettet ovenfor kan antas å komme fra en log-logistisk fordeling.

Forklar også kort hvordan estimater for μ og σ kan finnes fra plottet.

c) Finn den kumulative hasardfunksjon $Z(t)$ for T .

Hva er fordelingen for den tilfeldige variabelen $Z(T)$? (Du trenger ikke å bevise dette).

Forklar hvordan dette resultatet kan brukes til å sjekke om det høyresensurerte datasettet kan antas å komme fra en log-logistisk fordeling. Kjenner du navnet på denne metoden?

**Tabell over kumulative sannsynligheter
for χ^2 -fordelingen med 1 frihetsgrad.**

Chi-Square with 1 DF

x	P(X <= x)
0,1	0,248170
0,2	0,345279
0,3	0,416118
0,4	0,472911
0,5	0,520500
0,6	0,561422
0,7	0,597216
0,8	0,628907
0,9	0,657218
1,0	0,682689
1,1	0,705734
1,2	0,726678
1,3	0,745787
1,4	0,763276
1,5	0,779329
1,6	0,794097
1,7	0,807712
1,8	0,820288
1,9	0,831922
2,0	0,842701
2,1	0,852701
2,2	0,861989
2,3	0,870626
2,4	0,878665
2,5	0,886154
2,6	0,893136
2,7	0,899652
2,8	0,905736
2,9	0,911420
3,0	0,916735
3,1	0,921708
3,2	0,926362
3,3	0,930720
3,4	0,934804
3,5	0,938631
3,6	0,942220

Standard normalfordeling

$$\Phi(z) = P(Z \leq z)$$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
-3.7	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
-3.6	.0002	.0002	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001
-3.5	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002	.0002
-3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
-3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
-3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
-3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
-3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
-2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
-2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
-2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
-2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
-2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
-2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
-2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
-2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
-2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
-2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
-1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
-1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
-1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
-1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
-1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
-1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
-1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
-1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
-1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
-1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
-.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
-.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
-.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
-.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
-.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
-.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
-.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
-.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
-.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
-.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

Standard normalfordeling

$$\Phi(z) = P(Z \leq z)$$

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.9	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.0	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.1	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.2	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.3	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.4	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.5	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.6	.9452	.9463	.9474	.9484	.9495	.9505	.9515	.9525	.9535	.9545
1.7	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.8	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.9	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9761	.9767
2.0	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998
3.5	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998	.9998
3.6	.9998	.9998	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999
3.7	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999	.9999

Kritiske verdier i standard normalfordelingen

$$P(Z > z_\alpha) = \alpha$$

α	z_α
.2	0.842
.15	1.036
.1	1.282
.075	1.440
.05	1.645
.04	1.751
.03	1.881
.025	1.960
.02	2.054
.01	2.326
.005	2.576
.001	3.090
.0005	3.291
.0001	3.719
.00005	3.891
.00001	4.265
.000005	4.417
.000001	4.753