Norges teknisk-naturvitenskapelige universitet
Institutt for matematiske fag

Faglig kontakt under eksamen
Stian Lydersen tlf 73867270 / 92632393

Eksamen ST2303 Medisinsk statistikk
Torsdag 30 november 2006 kl 0900-1300

Tillatte hjelpemidler: Godkjent kalkulator. Tabeller og formler i statistikk, Institutt for matematiske fag. 1 håndskrevet A4-ark (stemplet ark fra Institutt for matematiske fag)

Eksamenssettet består av 6 sider norsk versjon, 6 sider engelsk versjon.

Oppgave 1

I et dobbelt blindt, randomisert klinisk forsøk sammenliknes to behandlinger A og B for en bestemt sykdom. 2 av 60 pasienter som ble behandlet med A ble friske, og 9 av 60 pasienter som ble behandlet med B, ble friske.

a)
Estimer sannsynlighetene p_A og p_B for å bli frisk hhv ved behandling A og behandling B. Beregn et 95% konfidensintervall for p_A - p_B.

Vi ønsker å teste om det er forskjell på behandlingene, dvs H_0: $p_A = p_B$ versus $p_A \neq p_B$.

b)
Fisher's eksakte test gir p-verdi=0.053. Forklar (med ord, ved å skrive opp en formel, og/eller ved å skissere en figur) hvordan verdien er regnet ut. Du skal ikke gjennomføre selve beregningen.

c)
Regn ut p-verdien for Pearson's kjikvadrattest uten Yates' kontinuitetskorreksjon. (Du kan bruke at dersom χ_1^2 er kjikvardratfordelt med 1 frihetsgrad, så er $P(\chi_1^2 \geq x) = 2P(Z \geq \sqrt{x})$ hvor Z er standard normalfordelt.) Er forutsetningene for testen oppfylt her?

d)
Hvilke gode og/eller dårlige egenskaper, har hhv Fisher's eksakte test og Pearson's kjikvardrattest? Hvilken test vil du anbefale her og hvorfor? Sammenlikne med svarene i a), b) og c) og kommenter.

Oppgave 2

Tabellen nedenfor viser klassifisering gjort av en radiolog for CT (computer tomografi) bilder for 99 pasienter der en hadde indikasjoner på nevrologiske problemer. Den sanne sykdoms-statusen ble også fastslått.

Count

		CT klassifiserinig av radiolog				
		sikkert frisk (1)	sannsynligvis frisk (2)	sannsynligvis syk (3)	sikkert syk (4)	Total
status	frisk (0)	31	6	10	2	49
	syk (1)	4	2	11	33	50
Total		35	8	21	35	99

a)
Anta i dette punktet at pasienten klassifiseres som ”syk” dersom radiologen klassifiserer vedkommende som ”sannsynligvis syk” (3) eller ”sikkert syk” (4). Estimer sensitivitet og spesifisitet i denne situasjonen.

b)
Tegn den empiriske ROC-kurven for datamaterialet. Ved øyemål, hva vil du anslå arealet under ROC-kurven til å være? Hvilke(n) fortolkning(er) har dette arealet? Hva vil arealet være hhv for en perfekt diagnostisk test, og for en fullstendig ikke-informativ diagnostisk test?

c)
I en logistisk regresjonsmodell med status som avhengig variabel og klassifisering som uavhengig variabel fås estimatene $\hat{\beta}_0 = -3.890$ og $\hat{\beta}_1 = 1.498$. Basert på disse tallene, estimer sannsynligheten for at en pasient som er klassifisert som ”sannsynligvis syk” (3), virkelig er syk. Virker svaret rimelig, hvis du sammenlikner med tabellen og/eller resultater ovenfor?

Oppgave 3

Multippel imputasjon (MI) er en metode for å håndtere manglende data.

a)
Forklart kort (f.eks med 4-5 setninger) hva MI går ut på.

b)
La Q være en populasjons-størrelse vi ønsker å estimere. La $\hat{Q}^{(j)}$ og $U^{(j)}$ være estimat og estimert varians for estimatoren, fra imputasjon nr j. Rubin’s formler for å kombinere resultatene er gitt ved

$$\bar{Q} = \frac{1}{m}\sum_{j=1}^{m}\hat{Q}^{(j)} \tag{1}$$

$$\bar{U} = \frac{1}{m}\sum_{j=1}^{m} U^{(j)} \qquad (2)$$

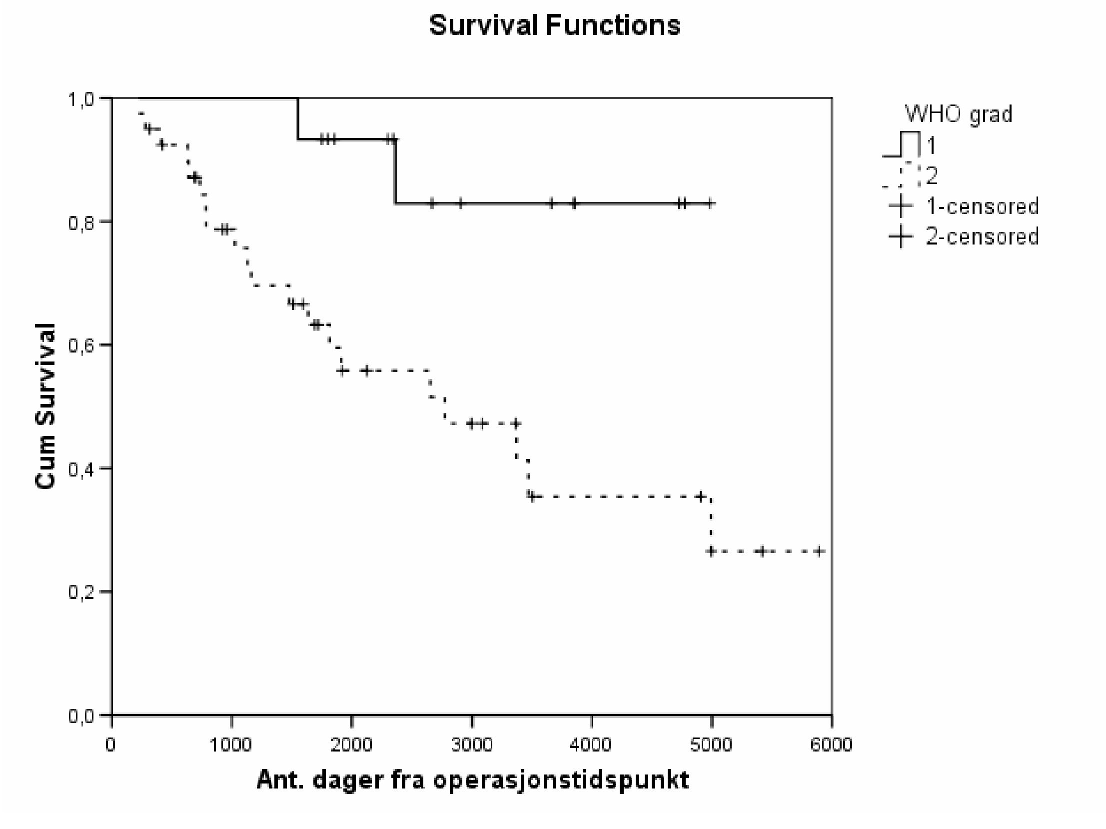
$$B = \frac{1}{m-1}\sum_{j=1}^{m}\left[\hat{Q}^{(j)} - \bar{Q}\right]^2 \qquad (3)$$

$$T = \bar{U} + \left(1 + \frac{1}{m}\right)B \qquad (4)$$

Hva regnes ut i hver av disse formlene? Vanligvis er verdien m=10 tiltrekkelig stor. Forklar kort hvorfor man ville oppnå lite ved å øke m særlig ut over dette.

Oppgave 4

Figur 1 viser Kaplan-Meier plott for overelevelse for 55 pasienter operert for hjernesvulst. Svulstene var gradert etter WHO grad 1 (15 pasienter) eller 2 (40 pasienter), hvor grad 2 er den mest aggressive. Pasientene var i alder 15-70 år ved operasjonstidspunktet, 33 av dem var menn. Observasjonstiden pr pasient varierte fra 200 til 5900 dager, og i denne tiden døde 2 av grad 1 pasientene, og 20 av grad 2 pasientene.



Figur 1

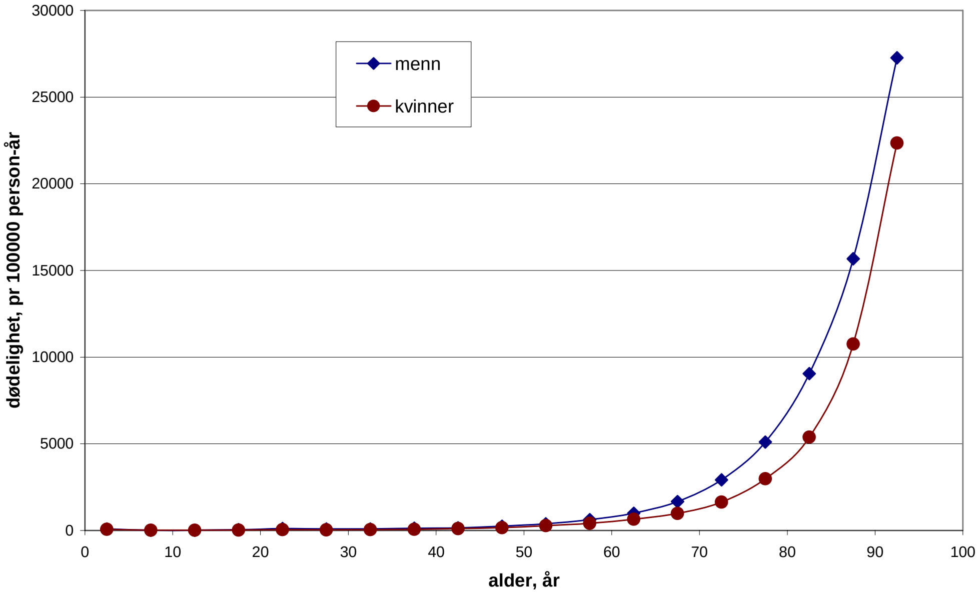
a)
Hva vil det si at en levetid er sensurert? Hvor mange av disse levetidene var sensurert? Basert på figuren, hva kan du si om median overlevelse for hver av de to WHO grad gruppene?

b)
Tabellen nedenfor viser resultater fra en Cox-analyse (Proportional Hazards Analysis) for det samme pasientmaterialet. Alder er gitt i år, og kjønn er kodet som mann (1) og kvinne (2).

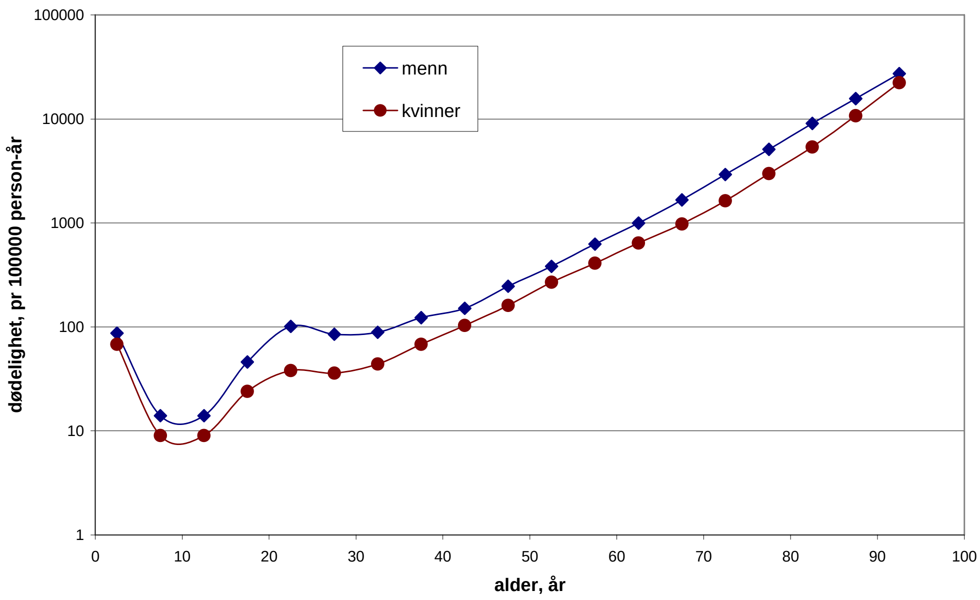
	B	SE
whograd	1,585	,748
alder	,043	,020
kjønn	-,632	,468

Regn ut p-verdien for WHO grad. Sammenlikne denne med P-verdien for log rang testen for WHO grad, som i dette tilfelle er 0.010. Hva forteller hver av de to p-verdiene?

c)
I mange Cox-analyser for tid til død inkluderer man alder og kjønn slik som ovenfor. Figur 2 og 3 viser dødelighet som funksjon av alder og kjønn i Norge, basert på tall fra Statistisk Sentralbyrå. (Bildet er liknende for andre land i vår del av verden.) Basert på Figur 2 og/eller Figur 3, kommenter om det er i samsvar med forutsetningene for Cox-modellen å inkludere alder og kjønn slik som ovenfor.



Figur 2: Dødelighet i Norge 2005, lineær Y-akse.



Figur 3: Dødelighet i Norge 2005, logaritmisk Y-akse.

Oppgave 5

Aktivert protein-C (APC) resitans kan være en risikofaktor for trombose (blodpropp). Tabellen nedenfor viser verdier målt på 10 pasienter, med 2 uavhengige målinger pr pasient (Tabell 2.17 i Rosner: Fundamentals of Biostatistics)

Case Summaries[a]

	id	APC
1	1	2,22
2	2	3,42
3	3	3,68
4	4	2,64
5	5	2,68
6	6	3,29
7	7	3,85
8	8	2,24
9	9	3,25
10	10	3,30
11	1	1,88
12	2	3,59
13	3	3,01
14	4	2,37
15	5	2,26
16	6	3,04
17	7	3,57
18	8	2,29
19	9	3,39
20	10	3,16
Total N	20	20

a. Limited to first 100 cases.

La Y_{ij} betegne j'te måling på pasient nr i. Anta at observasjonene er normaltfordelte. Skriv opp modellen for fordelingen til Y_{ij}. Skriv opp korrelasjonen mellom to målinger på samme pasient, uttrykt ved parametrene i denne modellen. Hva kalles denne korrelasjonen, og hvilken annen fortolkning har den enn en korrelasjon? Hva blir estimatet for denne, basert på (hele eller deler av) utskriften nedenfor?

Descriptive Statistics

Dependent Variable: APC

id	Mean	Std. Deviation	N
1	2,0500	,24042	2
2	3,5050	,12021	2
3	3,3450	,47376	2
4	2,5050	,19092	2
5	2,4700	,29698	2
6	3,1650	,17678	2
7	3,7100	,19799	2
8	2,2650	,03536	2
9	3,3200	,09899	2
10	3,2300	,09899	2
Total	2,9565	,58640	20

Variance Estimates

Component	Estimate
Var(id)	,309
Var(Error)	,051

Dependent Variable: APC

The Norwegian University of Science and Technology
Department of Mathematical Sciences

Academic contact during the exam:
Stian Lydersen, tel 73867270 / 92632393

Exam ST2303 Medical statistics
Thursday 30 November 2006, 0900-1300

Permitted aids: Approved calculator, Tabeller og formler i statistikk, Institutt for matematiske fag. 1 hand written A4 sheet (stamped sheet from Institutt for matematiske fag)

The exam consists of 6 pages Norwegian version, 6 pages English version.

Problem 1

Two treatments A and B for a certain disease were compared in a double blind, randomized clinical trial. 2 of 60 patients treated with A were cured, and 9 of 60 patients treated with B were cured.

a)
Estimate the probabilities p_A and p_B of cure for treatment A and for treatment B. Compute a 95% confidence interval for $p_A - p_B$.

We want to test if there is a difference between the treatments, that is, H_0: $p_A = p_B$ versus $p_A \neq p_B$.

b)
Fisher's exact test gives p-value= 0.053. Explain using words, a formula, and/or a figure how the value is computed. Do not carry out the computation of the value.

c)
Compute the p-value for Pearson's chi square test without Yates continuity correction. (You may use the fact that if χ_1^2 is chi square distributed with 1 degree of freedom, then $P(\chi_1^2 \geq x) = 2P(Z \geq \sqrt{x})$ where Z is standard normal distributed.) Are the assumptions for the test fulfilled here?

d)
What are the good and/or bad properties of Fisher's exact test and Pearson's chi square test, respectively? Which test would you recommend here, and why? Compare the answers in a), b), and c) and comment.

Problem 2

The table below shows a classification by a radiologist for CT (computer tomography) scans of 99 patients with indications of neurological problems. The true illness status was also recorded.

Count

		CT klassiferinig av radiolog				
		sikkert frisk (1)	sannsynligvis frisk (2)	sannsynligvis syk (3)	sikkert syk (4)	Total
status	frisk (0)	31	6	10	2	49
	syk (1)	4	2	11	33	50
Total		35	8	21	35	99

a)
Assume (here in point a) that the patient is classified as ”ill” if the radiologist classifies the patient as ”sannsynligvis syk” (3) or ”sikkert syk” (4). Estimate the sensitivity and specificity in this situation.

b)
Draw the empirical ROC-curve for the data set. Using a rough eye measure, what is the approximate area under the ROC-curve. Which interpretation does this area have? What is the area for a perfect test and for a completely non-informative test?

c)
In a logistic regresssion model with ”status” as dependent variable and ”klassifisering” as independent variable, we obtain the estimates $\hat{\beta}_0 = -3.890$ and $\hat{\beta}_1 = 1.498$. Based on these numbers, estimate the probability that a patient classified as”sannsynligvis syk” (3), really is ill. Is the answer reasonable, compared to the table and/or results above?

Oppgave 3

Multiple imputation (MI) is a method for handling missing data.

a)
Explain short (for example 4-5 sentences) the principle of MI.

b)
Let Q be a population quantity we want to estimate. Let $\hat{Q}^{(j)}$ and $U^{(j)}$ be its estimate and the estimated variance for the estimator, from imputation no j. Rubin’s formulas for combining the results are given by

$$\bar{Q} = \frac{1}{m}\sum_{j=1}^{m} \hat{Q}^{(j)} \tag{5}$$

$$\bar{U} = \frac{1}{m}\sum_{j=1}^{m} U^{(j)} \tag{6}$$

$$B = \frac{1}{m-1}\sum_{j=1}^{m}\left[\hat{Q}^{(j)} - \bar{Q}\right]^2 \tag{7}$$

$$T = \bar{U} + \left(1 + \frac{1}{m}\right)B \tag{8}$$

What is computed in each of these formulas? Usually the value m=10 is sufficiently large. Explain short why little would be obtained by increasing m much more than this.

Problem 4

Figure 1 shows a Kaplan-Meier plot for survival for 55 pasienter operated for brain tumor. The tumors were graded according to WHO grade 1 (15 patients) or 2 (40 patients), where grade 2 is the most aggressive. The patients were 15-70 years old at the time of operation, and 33 of them were men. The observation time per patient varied from 200 to 5900 days, and during the observation time 2 of the grade 1 pasients died, and 20 of the grade 2 pasients died.

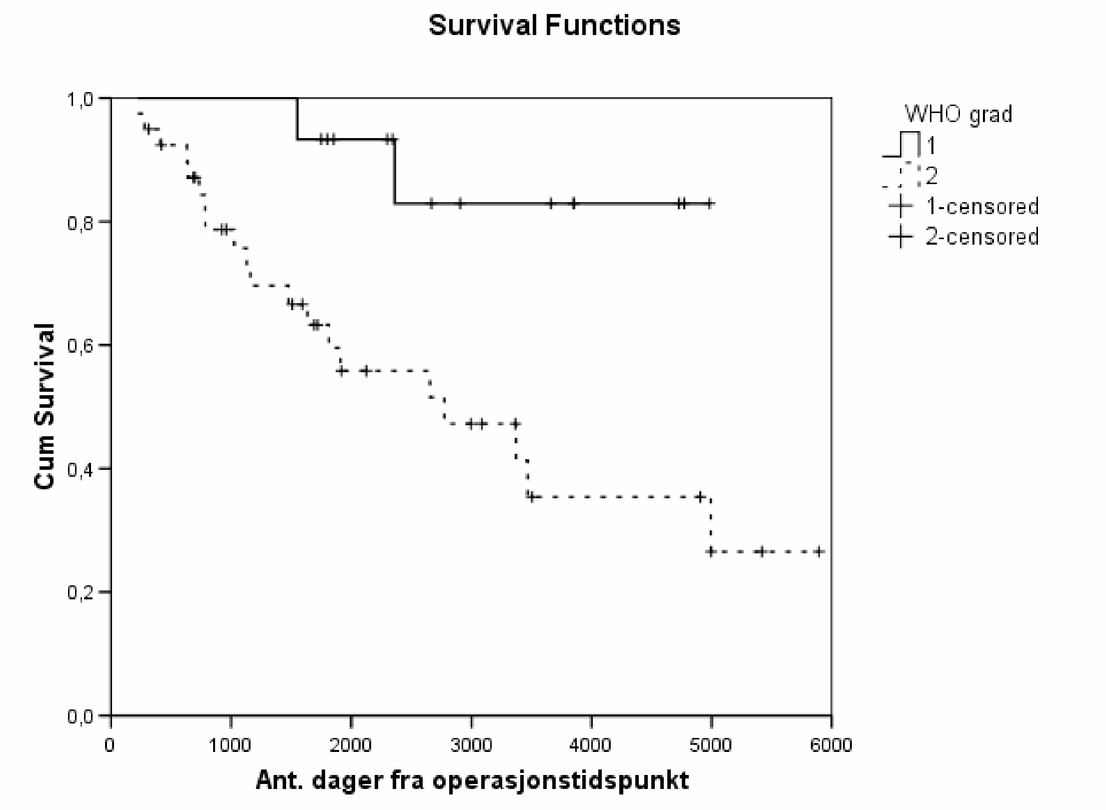


Figure 1. Survival probability as function of days from operation.

a)
What does it mean that a lifetime is censored? How many of these lifetimes were censored? Based on the figure, what can you say about the median survival for each of the two WHO grade groups?

b)
The table below shows results of a Cox (Proportional Hazards) analysis for the same patient material. Age ("alder") is given in years, and sex ("kjønn") is coded as man (1) and woman (2).

	B	SE
whograd	1,585	,748
alder	,043	,020
kjønn	-,632	,468

Compute the P-value for WHO grade. Compare it to the P-value for the log rank test for WHO grade, which is 0.010. What do the two P-values tell?

c)
In many Cox-analyses for time to death, age and sex is included in the way above. Figure 2 and 3 show mortality as function of age and sex in Norway, based on numbers from Statistisk Sentralbyrå. (The picture is similar for other countries in our part of the world.) Based on Figure 2 and/or Figure 3, comment whether including age and sex as above agrees with the assumptions for the Cox model.

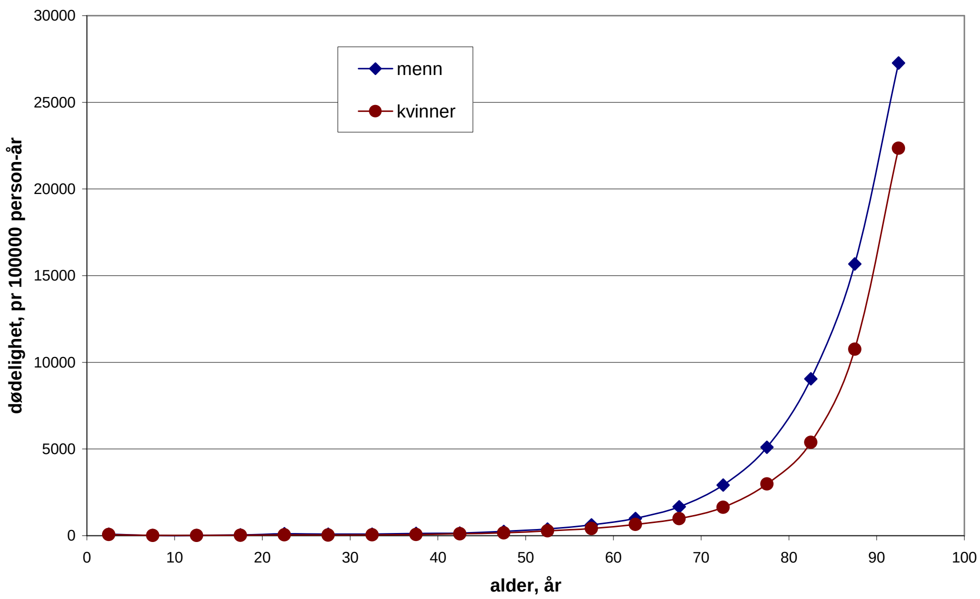


Figure 2: Mortality in Norway 2005, linear Y-axis.

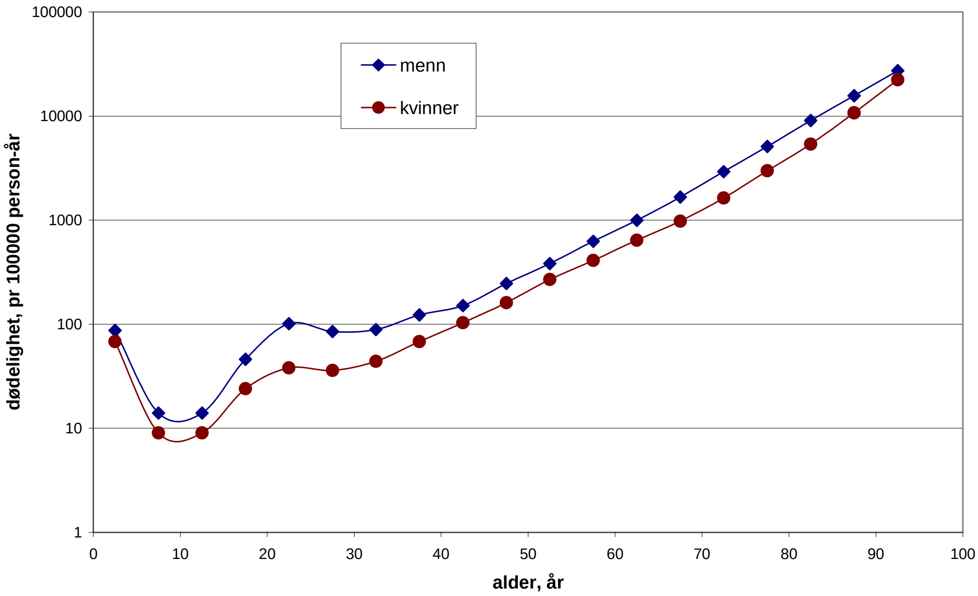


Figure 3: Mortality in Norway 2005, logarithmic Y-axis.

Problem 5

Activated protein-C (APC) resitance may be a risk factor for trombosis (blood clot). The table below shows values measured on 10 patients, with 2 independent measurements per patient (Table 2.17 in Rosner: Fundamentals of Biostatistics)

Case Summaries[a]

		id	APC
1		1	2,22
2		2	3,42
3		3	3,68
4		4	2,64
5		5	2,68
6		6	3,29
7		7	3,85
8		8	2,24
9		9	3,25
10		10	3,30
11		1	1,88
12		2	3,59
13		3	3,01
14		4	2,37
15		5	2,26
16		6	3,04
17		7	3,57
18		8	2,29
19		9	3,39
20		10	3,16
Total	N	20	20

a. Limited to first 100 cases.

LetY_{ij} denote measurement number j on patient number i. Assume normal distributed data. Write down the model for the distribution of Y_{ij}. Express the correlation between two measurements on the same patient, expressed by the parameters in the model. What is this correlation called, and which other interpretation than a correlation does it have? What is its estimate, based on (all or parts of) the printout below?

Descriptive Statistics

Dependent Variable: APC

id	Mean	Std. Deviation	N
1	2,0500	,24042	2
2	3,5050	,12021	2
3	3,3450	,47376	2
4	2,5050	,19092	2
5	2,4700	,29698	2
6	3,1650	,17678	2
7	3,7100	,19799	2
8	2,2650	,03536	2
9	3,3200	,09899	2
10	3,2300	,09899	2
Total	2,9565	,58640	20

Variance Estimates

Component	Estimate
Var(id)	,309
Var(Error)	,051

Dependent Variable: APC