Tentamen
MVE301 Sannolikhet, statistik och risk

2018-10-12 kl. 8:30 - 13:30

Examinator: Johan Jonasson, Matematiska vetenskaper, Chalmers

Telefonvakt: Olof Elias, telefon: 031-7725325

Hjälpmedel: Valfri miniräknare. Två blad (dvs fyra sidor) handskrivna anteckningar. Tabeller finns längst bak på tentamenstesen.

Denna tentamen utgör grunden för betygssättning. För betyg 3 krävs minst 18 poäng, för betyg 4 minst 27 poäng och för betyg 5 minst 36 poäng.

---

1. (6p) Från en normalfördelning med väntevärde och varians μ och σ^2 har det dragits ett stickprov:

$$\mathbf{X} = (-0.11,\ -0.72,\ 0.81,\ -0.26,\ -0.53,\ 1.83,\ 0.08,\ 0.36)$$

(a) Gör ett symmetriskt konfidensintervall av konfidensgrad 90 % under antagandet att σ^2 är okänd.

(b) Gör nu samma sak men antag nu att det är känt att $\sigma^2 = 1$.

2. (a) (4p) För data (x_k, Y_k), $k = 1, 2, 3, 4$ vill vi göra linjär regression enligt standardmodellen $Y_k = a + bx_k + \epsilon_k$, där ϵ_k:na är oberoende och $N(0, \sigma^2)$-fördelade. De data vi fick var $(20, 365)$, $(40, 220)$, $(60, 117)$ och $(80, 34)$. Skatta regressionslinjen $y = a + bx$ och gör ett 90% konfidensintervall för b. Kom ihåg att

$$s^2 = \frac{1}{n-2}\left(S_{yy} - \frac{S_{xy}^2}{S_{xx}}\right).$$

(b) (2p) Antag att man gjort en linjär regression och fått den skattade regressionslinjen $y = \hat{a} + \hat{b}x$ för data (x_k, y_k), $k = 1, \ldots, n$, där $x_1 < x_2 < \ldots < x_n$. Sedan visar det sig att man läst fel på sista y-värdet och att y_n istället skulle vara y_n' där $y_n' > y_n$. Man får då den nya skattade regressionslinjen $y = \hat{a}' + \hat{b}'x$. Visa att $\hat{b}' > \hat{b}$.

3. (6p) Låt T vara en stokastisk variabel med täthetsfunktion

$$f_T(t) = C(1 - t^2),\ t \in (0, 1).$$

Låt nu $X = T^2$ och $Y = \cos(\pi T)$. Bestäm konstanten C, beräkna väntevärdena för X och Y och $\mathbb{C}\text{ov}(X, Y)$.

4. (5p) I en hatt ligger tre mynt, mynt 1, 2 och 3, som är sådana att de vid slantsingling ger klave med sannolikhet 1/10, 1/20 respektive 1/30. Ett av mynten väljs på måfå och singlas n gånger och det visar sig att det inte blir någon klave alls. Bestäm den betingade sannolikheten givet detta, att det valda myntet var mynt i, $i = 1, 2, 3$.

5. (6p) En ishockeymatch spelas i tre perioder om vardera 20 minuter. Antag att målen i period 1 för respektive lag kommer som oberoende Poissonprocesser med parametrar 2 mål per timme för hemmalaget och 3/2 mål per timme för bortalaget. I period 2 är motsvarande siffror 3 mål per timme respektive 5/2 mål per timme och i period 3 är siffrorna 5 mål per timme respektive 4 mål per timme. Vad är förväntat totalt antal mål i matchen? Vad är sannolikheten att det blir totalt 6 mål? Vad är den betingade sannolikheten att hemmalaget vinner givet att det blir 6 mål?

(I en ishockeymatch blir det ibland förlängning för att utse en vinnare om resultatet är oavgjort vid full tid. Vi bortser från detta här och avser bara vad som sker under ordinarie tid.)

6. (5p) Låt $X_1, X_2, \ldots, X_n$ vara ett stickprov på en stokastisk variabel med täthetsfunktion

$$f_\theta(x) = \frac{2}{3}(x-\theta)^2, \; x \in (\theta - 1, \theta + 1),$$

där θ en okänd parameter. Visa att ML-skattningen $\hat{\theta}$ av θ uppfyller

$$\sum_{i=1}^{n} \frac{1}{x - \hat{\theta}} = 0.$$

Beräkna också $\hat{\theta}$ då $n = 3$ och observationerna var 1/4, 1/2 och 3/2.

7. (5p) Betrakta en Bayesiansk modell där θ_1 och θ_2 är två paramatrer som àpriori väljs oberoende och båda enligt en $\beta(2,1)$-fördelning. Låt sedan, givet θ_1 och θ_2, de stokastiska variablerna $X = (X_1, \ldots X_n)$ vara oberoende och sådana att $\mathbb{P}(X_i = 1) = 1 - \mathbb{P}(X_i = 0) = \theta_1\theta_2$.

Beräkna, sånär som på en proportionalitetskontant, àposterioritätheten $f_{\theta_1,\theta_2|X}(t_1, t_2|x)$ för (θ_1, θ_2) givet ett utfall $x = (x_1, \ldots, x_n)$ sådant att $\sum_{i=1}^n x_i = k$. Ange också MAP $(\hat{\theta}_1, \hat{\theta}_2)$ av (θ_1, θ_2), dvs de värden av (t_1, t_2) som maximerar denna àposterioritäthet.

8. (6p) Simulering av exponentialfördelning och Poissonfördelning utgående från likformig fördelning:

(a) Låt X vara en stokastisk variabel som är likformigt fördelad på $(0, 1)$. Bestäm en funktion g sådan att $g(X)$ är exponentialfördelad med parameter λ.

(b) Låt $T_1, T_2, \ldots$ vara oberoende och exponentialfördelade med parameter 1. Bestäm en funktion $h(T_1, T_2, \ldots)$ sådan att $h(T_1, T_2, \ldots)$ är Poissonfördelad med parameter λ.

Lycka till!
Johan Jonasson

Tabell 1: Values of the cdf $\Phi(x)$ of the standard normal distribution [e.g., $\Phi(1.41) = 0.921$]

x	0	1	2	3	4	5	6	7	8	9
0.0	.500	.504	.508	.512	.516	.520	.524	.528	.532	.536
0.1	.540	.544	.548	.552	.556	.560	.564	.568	.571	.575
0.2	.579	.583	.587	.591	.595	.599	.603	.606	.610	.614
0.3	.618	.622	.626	.629	.633	.637	.641	.644	.648	.652
0.4	.655	.659	.663	.666	.670	.674	.677	.681	.684	.688
0.5	.692	.695	.698	.702	.705	.709	.712	.716	.719	.722
0.6	.726	.729	.732	.736	.739	.742	.745	.749	.752	.755
0.7	.758	.761	.764	.767	.770	.773	.776	.779	.782	.785
0.8	.788	.791	.794	.797	.800	.802	.805	.808	.811	.813
0.9	.816	.819	.821	.824	.826	.829	.832	.834	.836	.839
1.0	.841	.844	.846	.848	.851	.853	.855	.858	.860	.862
1.1	.864	.867	.869	.871	.873	.875	.877	.879	.881	.883
1.2	.885	.887	.889	.891	.892	.894	.896	.898	.900	.902
1.3	.903	.905	.907	.908	.910	.912	.913	.915	.916	.918
1.4	.919	.921	.922	.924	.925	.926	.928	.929	.931	.932
1.5	.933	.934	.936	.937	.938	.939	.941	.942	.943	.944
1.6	.945	.946	.947	.948	.950	.951	.952	.952	.9545	.954
1.7	.955	.956	.957	.958	.959	.960	.961	.962	.962	.963
1.8	.964	.965	.966	.966	.967	.968	.969	.969	.970	.971
1.9	.971	.972	.973	.973	.974	.974	.975	.976	.976	.977
2.0	.977	.978	.978	.979	.979	.980	.980	.981	.981	.982
2.1	.982	.983	.983	.983	.984	.984	.985	.985	.985	.986
2.2	.986	.986	.987	.987	.988	.988	.988	.988	.989	.989
2.3	.989	.990	.990	.990	.990	.991	.991	.991	.991	.992
2.4	.992	.992	.992	.992	.993	.993	.993	.993	.993	.994
2.5	.994	.994	.994	.994	.995	.995	.995	.995	.995	.995
2.6	.995	.996	.996	.996	.996	.996	.996	.996	.996	.996
2.7	.996	.997	.997	.997	.997	.997	.997	.997	.997	.997
2.8	.997	.998	.998	.998	.998	.998	.998	.998	.998	.998
2.9	.998	.998	.998	.998	.998	.998	.998	.998	.999	.999

Tabell 2: Values of $\Phi(x)$ commonly used in confidence intervals and tests, and the corresponding x values

$\Phi(x)$	0.90	0.95	0.975	0.99	0.995
x	1.28	1.64	1.96	2.33	2.58

Tabell 3: Percentiles of the t distribution with DF degrees of freedom [e.g., $F_{t_7}(1.89) = 0.95$]

DF	0.95	0.975	0.99	0.995	DF	0.95	0.975	0.99	0.995
1	6.31	12.71	31.82	63.66	16	1.75	2.12	2.58	2.92
2	2.92	4.30	6.96	9.92	17	1.74	2.11	2.58	2.90
3	2.35	3.18	4,54	5.84	18	1.73	2.10	2.55	2.88
4	2.13	2.78	3.74	4.60	19	1.73	2.09	2.54	2.86
5	2.02	2.57	3.36	4.03	20	1.72	2.09	2.53	2.85
6	1.94	2.45	3.14	3.71	21	1.72	2.08	2.52	2.83
7	1.89	2.36	3.00	3.50	22	1.72	2.07	2.51	2.82
8	1.86	2.31	2.90	3.36	23	1.71	2.07	2.50	2.81
9	1.83	2.26	2.82	3.25	24	1.71	2.06	2.49	2.80
10	1.81	2.23	2.76	3.17	25	1.71	2.06	2.49	2.79
11	1.80	2.20	2.72	3.11	26	1.71	2.06	2.48	2.78
12	1.78	2.18	2.68	3.05	27	1.70	2.05	2.47	2.77
13	1.77	2.16	2.65	3.01	28	1.70	2.05	2.47	2.76
14	1.76	2.14	2.62	2.98	29	1.70	2.05	2.46	2.76
15	1.75	2.13	2.60	2.95	30	1.70	2.04	2.46	2.75

Tabell 4: Percentiles of the chi-square distribution with DF degrees of freedom [e.g., $F_{\chi^2_{20}}(10.85) = 0.05$]

DF	0.025	0.05	0.95	0.975	DF	0.025	0.05	0.95	0.975
1	0.001	0.004	3.84	5.02	16	6.91	7.96	26.30	28.84
2	0.05	0.10	5.99	7.38	17	7.56	8.67	27.59	30.19
3	0.22	0.35	7.82	9.34	18	8.23	9.39	28.87	31.53
4	0.48	0.71	9.49	11.14	19	8.91	10.12	30.14	32.85
5	0.83	1.14	11.07	12.83	20	9.59	10.85	31.41	34.17
6	1.24	1.64	12.59	14.45	21	10.28	11.60	32.67	35.48
7	1.69	2.17	14.07	16.01	22	10.98	12.34	33.92	36.78
8	2.18	2.73	15.51	17.54	23	11.69	13.09	35.17	38.08
9	2.70	3.32	19.92	19.02	24	12.40	13.85	36.42	39.36
10	3.25	3.94	18.31	20.48	25	13.12	14.61	37.65	40.65
11	3.82	4.58	19.68	21.92	26	13.84	15.38	38.88	41.92
12	4.40	5.23	21.03	23.34	27	14.57	16.15	40.11	43.19
13	5.01	5.89	22.36	27.74	28	15.31	16.93	41.34	44.46
14	5.63	6.57	23.68	26.12	29	16.05	17.71	42.56	45.72
15	6.26	7.26	25.00	27.49	30	16.79	18.49	43.77	46.98

Tabell 5: Percentiles of the F distribution with r and s degrees of freedom [e.g., $F_{F_{8,20}}(2.45) = 0.95$]

	2.5 % percentile								
s	$r = 2$	3	4	5	6	7	8	9	10
2	0.026	0.062	0.094	0.119	0.138	0.153	0.165	0.175	0.183
3	0.026	0.065	0.100	0.129	0.152	0.170	0.185	0.197	0.207
4	0.025	0.066	0.104	0.135	0.161	0.181	0.198	0.212	0.224
5	0.025	0.067	0.107	0.140	0.167	0.189	0.208	0.223	0.236
6	0.025	0.068	0.109	0.143	0.172	0.195	0.215	0.231	0.246
7	0.025	0.068	0.110	0.146	0.176	0.200	0.221	0.238	0.253
8	0.025	0.069	0.111	0.148	0.179	0.204	0.226	0.244	0.259
9	0.025	0.069	0.112	0.150	0.181	0.207	0.230	0.248	0.265
10	0.025	0.069	0.113	0.151	0.183	0.210	0.233	0.252	0.269
12	0.025	0.070	0.114	0.153	0.186	0.214	0.238	0.259	0.276
15	0.025	0.070	0.116	0.156	0.190	0.219	0.244	0.265	0.284
16	0.025	0.070	0.116	0.156	0.191	0.220	0.245	0.267	0.286
18	0.025	0.070	0.116	0.157	0.192	0.222	0.248	0.270	0.290
20	0.025	0.071	0.117	0.158	0.193	0.224	0.250	0.273	0.293
21	0.025	0.071	0.117	0.158	0.194	0.225	0.251	0.274	0.294
24	0.025	0.071	0.117	0.159	0.195	0.227	0.253	0.277	0.297
25	0.025	0.071	0.118	0.160	0.196	0.227	0.254	0.278	0.298
27	0.025	0.071	0.118	0.160	0.197	0.228	0.255	0.279	0.300
28	0.025	0.071	0.118	0.160	0.197	0.228	0.256	0.280	0.301
30	0.025	0.071	0.118	0.161	0.197	0.229	0.257	0.281	0.302

	95 % percentile								
s	$r = 2$	3	4	5	6	7	8	9	10
2	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40
3	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79
4	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96
5	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74
6	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06
7	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64
8	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35
9	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14
10	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98
12	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75
15	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54
16	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49
18	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41
20	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35
21	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32
24	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25
25	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24
27	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20
28	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19
30	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16

	97.5 % percentile								
s	$r = 2$	3	4	5	6	7	8	9	10
2	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39	39.40
3	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47	14.42
4	10.65	9.98	9.60	9.36	9.20	9.07	8.98	8.90	8.84
5	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68	6.62
6	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52	5.46
7	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82	4.76
8	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36	4.30
9	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03	3.96
10	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78	3.72
12	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44	3.37
15	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12	3.06
16	4.69	4.08	3.73	3.50	3.34	3.22	3.12	3.05	2.99
18	4.56	3.95	3.61	3.38	3.22	3.10	3.01	2.93	2.87
20	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84	2.77
21	4.42	3.82	3.48	3.25	3.09	2.97	2.87	2.80	2.73
24	4.32	3.72	3.38	3.15	2.99	2.87	2.78	2.70	2.64
25	4.29	3.69	3.35	3.13	2.97	2.85	2.75	2.68	2.61
27	4.24	3.65	3.31	3.08	2.92	2.80	2.71	2.63	2.57
28	4.22	3.63	3.29	3.06	2.90	2.78	2.69	2.61	2.55
30	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57	2.51

Tabell 6: Critical values c for the Wilcoxon signed rank test, where n is the sample size and $C = n(n+1) - c$ [e.g., if $n = 20$, then $P(W \leq 61) = P(W \geq 149) \approx 0.05$]

n	0.025	0.05	$n(n+1)/2$	n	0.025	0.05	$n(n+1)/2$
5	0	1	15	18	41	48	171
6	1	3	21	19	47	54	190
7	3	4	28	20	53	61	210
8	4	6	36	21	59	68	231
9	6	9	45	22	67	76	253
10	9	11	55	23	74	84	276
11	11	14	66	24	82	92	300
12	14	18	78	25	90	101	325
13	18	22	91	26	99	111	351
14	22	26	105	27	108	120	378
15	26	31	120	28	117	131	406
16	30	36	136	29	127	141	435
17	35	42	153	30	138	152	465

Tabell 7: Critical values c for the Wilcoxon rank sum test, where m is the size of the smaller sample, and $C = m(m+n+1) - c$ [e.g., if $m = 4$ and $n = 8$, then $P(W \leq 16) = P(W \geq 36) \approx 0.05$]

n	$P(W \leq c)$	$m = 2$	3	4	5	6	7	8	9	10	11
2	0.025	3									
	0.05	3									
3	0.025	3	3								
	0.05	6	7								
4	0.025	3	6	11							
	0.05	3	7	12							
5	0.025	3	7	12	18						
	0.05	4	8	13	20						
6	0.025	3	8	13	19	27					
	0.05	4	9	14	21	29					
7	0.025	3	8	14	21	28	37				
	0.05	4	9	15	22	30	40				
8	0.025	4	9	15	22	30	39	50			
	0.05	5	10	16	24	32	42	52			
9	0.025	4	9	15	23	32	41	52	63		
	0.05	5	11	17	25	34	44	55	67		
10	0.025	4	10	16	24	33	43	54	66	79	
	0.05	5	11	18	27	36	46	57	70	83	
11	0.025	5	10	17	25	35	45	56	69	82	97
	0.05	5	12	19	28	38	48	60	73	87	101

Tentamen
MVE301 Sannolikhet, statistik och risk

2018-10-12 kl. 8:30 - 13:30

Examinator: Johan Jonasson, Matematiska vetenskaper, Chalmers

Telefonvakt: Olof Elias, telefon: 031-7725325

Hjälpmedel: Valfri miniräknare. Två blad (dvs fyra sidor) handskrivna anteckningar. Tabeller finns längst bak på tentamenstesen.

Denna tentamen utgör grunden för betygssättning. För betyg 3 krävs minst 18 poäng, för betyg 4 minst 27 poäng och för betyg 5 minst 36 poäng.

---

1. (6p) Från en normalfördelning med väntevärde och varians μ och σ^2 har det dragits ett stickprov:

$$\mathbf{X} = (-0.11,\ -0.72,\ 0.81,\ -0.26,\ -0.53,\ 1.83,\ 0.08,\ 0.36)$$

(a) Gör ett symmetriskt konfidensintervall av konfidensgrad 90 % under antagandet att σ^2 är okänd.

(b) Gör nu samma sak men antag nu att det är känt att $\sigma^2 = 1$.

Lösning. Det här är standarduppgifter. För del (a) gäller att $(\overline{X} - \mu)/(s/\sqrt{n}) \sim t_{n-1}$ och det symmetriska konfidensintervallet för μ blir då, eftersom $n = 8$,

$$\mu\overline{X} \pm F_{t_7}^{-1}(0.95)\frac{s}{\sqrt{8}} = 0.185 \pm 0.552.$$

I (b) använder vi istället att $(\overline{X} - \mu)/(\sigma/\sqrt{n}) \sim N(0, 1)$, så

$$\mu\overline{X} \pm \Phi^{-1}(0.95)\frac{\sigma}{\sqrt{8}} = 0.185 \pm 0.580.$$

(Det kan vara lite överraskande att det blir ett kortare konfidensintervall med okänd varians, men konfidensintervall är själva slumpmässiga och i det här fallet råkade stickprovsvariansen bli mindre än 1 trots att den äkta variansen var 1.)

2. (a) (4p) För data (x_k, Y_k), $k = 1, 2, 3, 4$ vill vi göra linjär regression enligt standardmodellen $Y_k = a + bx_k + \epsilon_k$, där ϵ_k:na är oberoende och $N(0, \sigma^2)$-fördelade. De data vi fick var $(20, 365)$, $(40, 220)$, $(60, 117)$ och $(80, 34)$. Skatta regressionslinjen $y = a + bx$ och gör ett 90% konfidensintervall för b. Kom ihåg att

$$s^2 = \frac{1}{n-2}\left(S_{yy} - \frac{S_{xy}^2}{S_{xx}}\right).$$

(b) (2p) Antag att man gjort en linjär regression och fått den skattade regressionslinjen $y = \hat{a} + \hat{b}x$ för data (x_k, y_k), $k = 1, \ldots, n$, där $x_1 < x_2 < \ldots < x_n$. Sedan visar det sig att man läst fel på sista y-värdet och att y_n istället skulle vara y'_n där $y'_n > y_n$. Man får då den nya skattade regressionslinjen $y = \hat{a}' + \hat{b}'x$. Visa att $\hat{b}' > \hat{b}$.

Lösning: (a) Vi behöver S_{xx}, S_{xy} och S_{yy}. Några snabba uträkningar ger

$$\sum x_k = 200,\ \sum x_k^2 = 12000,\ \sum y_k = 736,\ \sum y_k^2 = 196470,\ \sum x_k y_k = 25840.$$

Enligt standardformel fås

$$S_{xy} = \sum x_k y_k - \frac{1}{4} \sum x_k \sum y_k = -10960$$

och som specialfall av denna formel

$$S_{xx} = 2000, \; S_{yy} = 61046.$$

Vi får då

$$s^2 = \frac{1}{2} \left(S_{yy} - \frac{S_{xy}^2}{S_{xx}} \right) = 492.6$$

så $s = \sqrt{492.6} \approx 22.195$. Nu fås

$$\hat{b} = \frac{S_{xy}}{S_{xx}} = -5.48$$

$$\hat{a} = \overline{y} - \hat{b}\overline{x} = \frac{1}{4}(\sum y_k - \hat{b} \sum x_k) = 458.$$

Den skattade regressionslinjen är alltså

$$y = 458 - 5.48x.$$

Konfidensintervallet ges av

$$b = \hat{b} \pm F_{t_2}^{-1}(0.95) \frac{s}{\sqrt{S_{xx}}} = -5.48 \pm 1.45 \; (90\%).$$

(b) Eftersom $\hat{b} = S_{xy}/S_{xx}$ och S_{xx} inte beror av y_k:na handlar det om att visa att S_{xy} är växande i y_n. För detta är det tillräckligt att visa att $\partial S_{xy}/\partial y_n > 0$. Men

$$S_{xy} = \sum (y_k - \frac{1}{n} \sum y_j)(x_k - \overline{x})$$

så

$$\frac{\partial S_{xy}}{\partial y_n} = -\frac{1}{n} \sum (x_k - \overline{x}) + (x_n - \overline{x}) = x_n - \overline{x} > 0.$$

3. (6p) Låt T vara en stokastisk variabel med täthetsfunktion

$$f_T(t) = C(1 - t^2), \; t \in (0, 1).$$

Låt nu $X = T^2$ och $Y = \cos(\pi T)$. Bestäm konstanten C, beräkna väntevärdena för X och Y och $\mathbb{C}\mathrm{ov}(X, Y)$.

Lösning. Konstanten bestäms av att

$$\int_0^1 C(1 - t^2) dt = C \left[t - \frac{t^3}{3} \right]_0^1 = C \frac{2}{3} = 1$$

vilket ger $C = 3/2$. Enligt den omedvetne statistikerns lag gäller

$$\mathbb{E}[X] = \mathbb{E}[T^2] = \int_0^1 t^2 f(t) dt = \frac{3}{2} \int_0^1 t^2 (1 - t^2) dt = \frac{3}{2} \left(\frac{1}{3} - \frac{1}{5} \right) = \frac{1}{5}$$

$$\mathbb{E}[Y] = \mathbb{E}[\cos(\pi T)] = \frac{3}{2} \int_0^1 (1 - t^2) \cos(\pi t) dt = \frac{3}{\pi^2}$$

och

$$\mathbb{E}[XY] = \mathbb{E}[T^2 \cos(\pi T)] = \frac{3}{2} \int_0^1 t^2 (1 - t^2) \cos(\pi t) dt = -\frac{12 - \pi^2}{\pi^4}$$

och därmed också

$$\mathbb{C}\mathrm{ov}(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = -\frac{3(12 - \pi^2)}{\pi^4} - \frac{3}{5\pi^2} = -\frac{12(15 - \pi^2)}{5\pi^4}.$$

4. (5p) I en hatt ligger tre mynt, mynt 1, 2 och 3, som är sådana att de vid slantsingling ger klave med sannolikhet 1/10, 1/20 respektive 1/30. Ett av mynten väljs på måfå och singlas n gånger och det visar sig att det inte blir någon klave alls. Bestäm den betingade sannolikheten givet detta, att det valda myntet var mynt i, $i = 1, 2, 3$.

Lösning. Låt A_i vara händelsen att mynt nummer i väljs, $i = 1, 2, 3$. Låt X vara antalet klave som de n slantsinglingarna resulterar i. Det som söks är $\mathbb{P}(A_i|X = 0)$. Enligt Bayes formel är detta proportionellt mot $\mathbb{P}(X = 0|A_i)\mathbb{P}(A_i)$ och eftersom $\mathbb{P}(A_i)$ i detta fall är densamma för alla i så blir detta i sin tur proportionellt mot $\mathbb{P}(X = 0|A_t)$ i det här fallet. Nu är ju, eftersom sannolikheten att få klave för mynt i är $1/(10i)$,

$$\mathbb{P}(X = 0|A_i) = \left(1 - \frac{1}{10i}\right)^n = \left(\frac{10i - 1}{10i}\right)^n.$$

För $i = 1, 2, 3$ blir dessa $(9/10)^n$, $(19/20)^n$ respektive $((29/30)^n$. För att slippa bråkuttryck kan vi om vi vill, skriv bråken i termer av sextiondelar och notera att dessa är proportionella mot 54^n, 57^n respektive 58^n. Alltså:

$$\mathbb{P}(A_1|X = 0) = \frac{54^n}{54^n + 57^n + 58^n}, \mathbb{P}(A_2|X = 0) = \frac{57^n}{54^n + 57^n + 58^n}$$

och

$$\mathbb{P}(A_3|X = 0) = \frac{58^n}{54^n + 57^n + 58^n}.$$

5. (6p) En ishockeymatch spelas i tre perioder om vardera 20 minuter. Antag att målen i period 1 för respektive lag kommer som oberoende Poissonprocesser med parametrar 2 mål per timme för hemmalaget och 3/2 mål per timme för bortalaget. I period 2 är motsvarande siffror 3 mål per timme respektive 5/2 mål per timme och i period 3 är siffrorna 5 mål per timme respektive 4 mål per timme. Vad är förväntat totalt antal mål i matchen? Vad är sannolikheten att det blir totalt 6 mål? Vad är den betingade sannolikheten att hemmalaget vinner givet att det blir 6 mål?

(I en ishockeymatch blir det ibland förlängning för att utse en vinnare om resultatet är oavgjort vid full tid. Vi bortser från detta här och avser bara vad som sker under ordinarie tid.)

Lösning. Antalet impulser i en Poissonprocess med intensitet λ i ett tidsintervall om längd t är Poissonfördelat med parameter λt. Låt H vara antalet mål som hemmalaget gör och B antalet mål som bortalaget gör. Etersom summor av oberoende Poissonfördelade stokastiska variabler är Poissonfördelad med en parameter som är summan av de ingående parametrarna och en period varar 1/3 timme, får vi

$$H \sim Poi\left(\frac{2 + 3 + 5}{3}\right) = Poi\left(\frac{10}{3}\right)$$

och på samma sätt att $B \sim Poi(8/3)$. Vi söker först $\mathbb{E}[H + B]$ som vi direkt ser att vara lika med $10/3 + 8/3 = 6$. Härnäst räknar vi ut

$$\mathbb{P}(H + B = 6) = e^{-6}\frac{6^6}{6!} = \frac{324}{5}e^{-6} \approx 0.161.$$

Till sist

$$\begin{aligned}
\mathbb{P}(H > B|H + B = 6) &= \frac{\mathbb{P}(H + B = 6, H > B)}{\mathbb{P}(H + B = 6)} \\
&= \frac{\mathbb{P}(H = 6, B = 0) + \mathbb{P}(H = 5, B = 1) + \mathbb{P}(H = 4, B = 2)}{\mathbb{P}(H + B = 6)} \\
&= \frac{\frac{(10/3)^6}{6!}\frac{(8/3)^0}{0!} + \frac{(10/3)^5}{5!}\frac{(8/3)}{1!} + \frac{(10/3)^4}{4!}\frac{(8/3)^2}{2!}}{\frac{6^6}{6!}} \\
&\approx 0.453.
\end{aligned}$$

6. (5p) Låt $X_1, X_2, \ldots, X_n$ vara ett stickprov på en stokastisk variabel med täthetsfunktion

$$f_\theta(x) = \frac{2}{3}(x-\theta)^2,\ x \in (\theta-1, \theta+1),$$

där θ en okänd parameter. Visa att ML-skattningen $\hat{\theta}$ av θ uppfyller

$$\sum_{i=1}^n \frac{1}{x-\hat{\theta}} = 0.$$

Beräkna också $\hat{\theta}$ då $n = 3$ och observationerna var 1/4, 1/2 och 3/2.

Lösning. Det gäller att

$$L(\theta) = L(\theta; x_1, \ldots, x_n) = \left(\frac{2}{3}\right)^n \prod_{i=1}^n (x_i - \theta)^2$$

då $x_i - 1 < \theta < x_i + 1$ för alla i och 0 annars. Alltså är $\hat{\theta}$ det $\theta \in (\max_i x_i - 1, \min_i x_i + 1)$ som maximerar det givna uttrycket. Det gäller att

$$l(\theta) = \log L(\theta) = n \ln\left(\frac{2}{3}\right) + 2\sum_{i=1}^n \ln|x_i - \theta|$$

och att ta $l'(\theta) = 0$ ger

$$\sum_{i=1}^n \frac{1}{x_i - \theta} = 0$$

som önskat. I det givna specialfallet ska vi lösa

$$\frac{1}{\theta - x_1} + \frac{1}{\theta - x_2} + \frac{1}{\theta - x_3} = 0$$

vilket ger andragradsekvationen

$$3\theta^2 - 2(x_1 + x_2 + x_3)\theta + x_1x_2 + x_1x_3 + x_2x_3 = 0$$

med lösningar

$$\theta = \overline{x} \pm \sqrt{\overline{x}^2 - \frac{1}{3}(x_1x_2 + x_1x_3 + x_2x_3)}$$

som med de givna siffrorna ger

$$\theta = \frac{9 \pm \sqrt{21}}{12}.$$

Den lösning som är aktuell är den som ligger i intervallet $(1/2, 5/4)$, dvs vi får

$$\hat{\theta} = \frac{9 + \sqrt{21}}{12} \approx 1.13.$$

7. (5p) Betrakta en Bayesiansk modell där θ_1 och θ_2 är två paramatrer som àpriori väljs oberoende och båda enligt en $\beta(2,1)$-fördelning. Låt sedan, givet θ_1 och θ_2, de stokastiska variablerna $X = (X_1, \ldots X_n)$ vara oberoende och sådana att $\mathbb{P}(X_i = 1) = 1 - \mathbb{P}(X_i = 0) = \theta_1\theta_2$.

Beräkna, sånär som på en proportionalitetskontant, àposterioritätheten $f_{\theta_1,\theta_2|X}(t_1, t_2|x)$ för (θ_1, θ_2) givet ett utfall $x = (x_1, \ldots, x_n)$ sådant att $\sum_{i=1}^n x_i = k$. Ange också MAP $(\hat{\theta}_1, \hat{\theta}_2)$ av (θ_1, θ_2), dvs de värden av (t_1, t_2) som maximerar denna àposterioritäthet.

Lösning. Vi har

$$\begin{aligned} f_{\theta_1,\theta_2|X}(t_1,t_2|x) &\propto f_{X|\theta_1,\theta_2}(x|t_1,t_2)f_{\theta_1,\theta_2}(t_1,t_2) \\ &= (t_1t_2)^k(1-t_1t_2)^{n-k}t_1^2t_2^2(1-t_1)(1-t_2) \\ &= t_1^{k+2}t_2^{k+2}(1-t_1)(1-t_2)(1-t_1t_2)^{n-k}. \end{aligned}$$

Av symmetrin mellan t_1 och t_2 förstår vi att maximum måste hittas då $t_1 = t_2 =: t$. Med den likheten blir högerledet

$$G(t) = t^{2(k+2)}(1-t)^2(1-t^2)^{n-k}.$$

Ta logaritmen, derivera och sätt till 0 och få

$$(k+2)(1-t^2) - t(1+t) - (n-k)t^2 = 0,$$

som ger

$$t = \frac{1+\sqrt{1+4(k+2)(n+3)}}{2(n+3)}.$$

8. (6p) Simulering av exponentialfördelning och Poissonfördelning utgående från likformig fördelning:

(a) Låt X vara en stokastisk variabel som är likformigt fördelad på $(0,1)$. Bestäm en funktion g sådan att $g(X)$ är exponentialfördelad med parameter λ.

(b) Låt $T_1, T_2, \ldots$ vara oberoende och exponentialfördelade med parameter 1. Bestäm en funktion $h(T_1, T_2, \ldots)$ sådan att $h(T_1, T_2, \ldots)$ är Poissonfördelad med parameter λ.

Lösning. Eftersom $X \sim likf(0,1)$ gäller att $F(x) = x$, $x \in (0,1)$. Om Y är en stokastisk variabel med fördelningsunktion F_Y och F_Y är inverterbar, gäller att

$$\mathbb{P}(F_Y^{-1}(X) \le x) = \mathbb{P}(X \le F_Y(x)) = F_Y(x),$$

dvs $F_Y^{-1}(X)$ har samma fördelning som Y. I (a) tillämpar vi detta med $Y \sim exp(\lambda)$, dvs $F_Y(x) = 1 - e^{-\lambda x}$. Genom att lösa ekvationen $y = 1 - e^{-\lambda x}$ m.a.p. x får vi att $F_Y^{-1}(y) = -\ln(1-y)/\lambda$. Denna funktion ger alltså ett svar till (a); vi kan ta

$$g(x) = -\frac{\ln(1-x)}{\lambda}.$$

För (b), kom ihåg att antal impulser i en Poissonprocess med intensitet 1 i tidsintervallet $[0,\lambda)$ är $Poi(\lambda)$-fördelad. Nu ges ju en Poissonprocess med intesitet 1 av att mellanrummen mellan impulser är oberoende och exponentialfördelade med intensitet 1. Med andra ord, om man skapar en process genom att låta impulser komma precis vid var och en av tidpunkterna T_1, $T_1 + T_2$, $T_1 + T_2 + T_2$ etc, får man en Poissonprocess med intensitet 1. Antalet impulser före tid λ blir då det största n sådant att $T_1 + \ldots + T_n < \lambda$. Ta alltså

$$h(T_1, T_2, \ldots) = \max\{n \in \mathbb{N} : \sum_{i=1}^{n} T_i < \lambda\}.$$

Lycka till!
Johan Jonasson

Tabell 1: Values of the cdf $\Phi(x)$ of the standard normal distribution [e.g., $\Phi(1.41) = 0.921$]

x	0	1	2	3	4	5	6	7	8	9
0.0	.500	.504	.508	.512	.516	.520	.524	.528	.532	.536
0.1	.540	.544	.548	.552	.556	.560	.564	.568	.571	.575
0.2	.579	.583	.587	.591	.595	.599	.603	.606	.610	.614
0.3	.618	.622	.626	.629	.633	.637	.641	.644	.648	.652
0.4	.655	.659	.663	.666	.670	.674	.677	.681	.684	.688
0.5	.692	.695	.698	.702	.705	.709	.712	.716	.719	.722
0.6	.726	.729	.732	.736	.739	.742	.745	.749	.752	.755
0.7	.758	.761	.764	.767	.770	.773	.776	.779	.782	.785
0.8	.788	.791	.794	.797	.800	.802	.805	.808	.811	.813
0.9	.816	.819	.821	.824	.826	.829	.832	.834	.836	.839
1.0	.841	.844	.846	.848	.851	.853	.855	.858	.860	.862
1.1	.864	.867	.869	.871	.873	.875	.877	.879	.881	.883
1.2	.885	.887	.889	.891	.892	.894	.896	.898	.900	.902
1.3	.903	.905	.907	.908	.910	.912	.913	.915	.916	.918
1.4	.919	.921	.922	.924	.925	.926	.928	.929	.931	.932
1.5	.933	.934	.936	.937	.938	.939	.941	.942	.943	.944
1.6	.945	.946	.947	.948	.950	.951	.952	.952	.9545	.954
1.7	.955	.956	.957	.958	.959	.960	.961	.962	.962	.963
1.8	.964	.965	.966	.966	.967	.968	.969	.969	.970	.971
1.9	.971	.972	.973	.973	.974	.974	.975	.976	.976	.977
2.0	.977	.978	.978	.979	.979	.980	.980	.981	.981	.982
2.1	.982	.983	.983	.983	.984	.984	.985	.985	.985	.986
2.2	.986	.986	.987	.987	.988	.988	.988	.988	.989	.989
2.3	.989	.990	.990	.990	.990	.991	.991	.991	.991	.992
2.4	.992	.992	.992	.992	.993	.993	.993	.993	.993	.994
2.5	.994	.994	.994	.994	.995	.995	.995	.995	.995	.995
2.6	.995	.996	.996	.996	.996	.996	.996	.996	.996	.996
2.7	.996	.997	.997	.997	.997	.997	.997	.997	.997	.997
2.8	.997	.998	.998	.998	.998	.998	.998	.998	.998	.998
2.9	.998	.998	.998	.998	.998	.998	.998	.998	.999	.999

Tabell 2: Values of $\Phi(x)$ commonly used in confidence intervals and tests, and the corresponding x values

$\Phi(x)$	0.90	0.95	0.975	0.99	0.995
x	1.28	1.64	1.96	2.33	2.58

Tabell 3: Percentiles of the t distribution with DF degrees of freedom [e.g., $F_{t_7}(1.89) = 0.95$]

DF	0.95	0.975	0.99	0.995	DF	0.95	0.975	0.99	0.995
1	6.31	12.71	31.82	63.66	16	1.75	2.12	2.58	2.92
2	2.92	4.30	6.96	9.92	17	1.74	2.11	2.58	2.90
3	2.35	3.18	4,54	5.84	18	1.73	2.10	2.55	2.88
4	2.13	2.78	3.74	4.60	19	1.73	2.09	2.54	2.86
5	2.02	2.57	3.36	4.03	20	1.72	2.09	2.53	2.85
6	1.94	2.45	3.14	3.71	21	1.72	2.08	2.52	2.83
7	1.89	2.36	3.00	3.50	22	1.72	2.07	2.51	2.82
8	1.86	2.31	2.90	3.36	23	1.71	2.07	2.50	2.81
9	1.83	2.26	2.82	3.25	24	1.71	2.06	2.49	2.80
10	1.81	2.23	2.76	3.17	25	1.71	2.06	2.49	2.79
11	1.80	2.20	2.72	3.11	26	1.71	2.06	2.48	2.78
12	1.78	2.18	2.68	3.05	27	1.70	2.05	2.47	2.77
13	1.77	2.16	2.65	3.01	28	1.70	2.05	2.47	2.76
14	1.76	2.14	2.62	2.98	29	1.70	2.05	2.46	2.76
15	1.75	2.13	2.60	2.95	30	1.70	2.04	2.46	2.75

Tabell 4: Percentiles of the chi-square distribution with DF degrees of freedom [e.g., $F_{\chi^2_{20}}(10.85) = 0.05$]

DF	0.025	0.05	0.95	0.975	DF	0.025	0.05	0.95	0.975
1	0.001	0.004	3.84	5.02	16	6.91	7.96	26.30	28.84
2	0.05	0.10	5.99	7.38	17	7.56	8.67	27.59	30.19
3	0.22	0.35	7.82	9.34	18	8.23	9.39	28.87	31.53
4	0.48	0.71	9.49	11.14	19	8.91	10.12	30.14	32.85
5	0.83	1.14	11.07	12.83	20	9.59	10.85	31.41	34.17
6	1.24	1.64	12.59	14.45	21	10.28	11.60	32.67	35.48
7	1.69	2.17	14.07	16.01	22	10.98	12.34	33.92	36.78
8	2.18	2.73	15.51	17.54	23	11.69	13.09	35.17	38.08
9	2.70	3.32	19.92	19.02	24	12.40	13.85	36.42	39.36
10	3.25	3.94	18.31	20.48	25	13.12	14.61	37.65	40.65
11	3.82	4.58	19.68	21.92	26	13.84	15.38	38.88	41.92
12	4.40	5.23	21.03	23.34	27	14.57	16.15	40.11	43.19
13	5.01	5.89	22.36	27.74	28	15.31	16.93	41.34	44.46
14	5.63	6.57	23.68	26.12	29	16.05	17.71	42.56	45.72
15	6.26	7.26	25.00	27.49	30	16.79	18.49	43.77	46.98

Tabell 5: Percentiles of the F distribution with r and s degrees of freedom [e.g., $F_{F_{8,20}}(2.45) = 0.95$]

	2.5 % percentile								
s	$r = 2$	3	4	5	6	7	8	9	10
2	0.026	0.062	0.094	0.119	0.138	0.153	0.165	0.175	0.183
3	0.026	0.065	0.100	0.129	0.152	0.170	0.185	0.197	0.207
4	0.025	0.066	0.104	0.135	0.161	0.181	0.198	0.212	0.224
5	0.025	0.067	0.107	0.140	0.167	0.189	0.208	0.223	0.236
6	0.025	0.068	0.109	0.143	0.172	0.195	0.215	0.231	0.246
7	0.025	0.068	0.110	0.146	0.176	0.200	0.221	0.238	0.253
8	0.025	0.069	0.111	0.148	0.179	0.204	0.226	0.244	0.259
9	0.025	0.069	0.112	0.150	0.181	0.207	0.230	0.248	0.265
10	0.025	0.069	0.113	0.151	0.183	0.210	0.233	0.252	0.269
12	0.025	0.070	0.114	0.153	0.186	0.214	0.238	0.259	0.276
15	0.025	0.070	0.116	0.156	0.190	0.219	0.244	0.265	0.284
16	0.025	0.070	0.116	0.156	0.191	0.220	0.245	0.267	0.286
18	0.025	0.070	0.116	0.157	0.192	0.222	0.248	0.270	0.290
20	0.025	0.071	0.117	0.158	0.193	0.224	0.250	0.273	0.293
21	0.025	0.071	0.117	0.158	0.194	0.225	0.251	0.274	0.294
24	0.025	0.071	0.117	0.159	0.195	0.227	0.253	0.277	0.297
25	0.025	0.071	0.118	0.160	0.196	0.227	0.254	0.278	0.298
27	0.025	0.071	0.118	0.160	0.197	0.228	0.255	0.279	0.300
28	0.025	0.071	0.118	0.160	0.197	0.228	0.256	0.280	0.301
30	0.025	0.071	0.118	0.161	0.197	0.229	0.257	0.281	0.302

	95 % percentile								
s	$r = 2$	3	4	5	6	7	8	9	10
2	19.00	19.16	19.25	19.30	19.33	19.35	19.37	19.38	19.40
3	9.55	9.28	9.12	9.01	8.94	8.89	8.85	8.81	8.79
4	6.94	6.59	6.39	6.26	6.16	6.09	6.04	6.00	5.96
5	5.79	5.41	5.19	5.05	4.95	4.88	4.82	4.77	4.74
6	5.14	4.76	4.53	4.39	4.28	4.21	4.15	4.10	4.06
7	4.74	4.35	4.12	3.97	3.87	3.79	3.73	3.68	3.64
8	4.46	4.07	3.84	3.69	3.58	3.50	3.44	3.39	3.35
9	4.26	3.86	3.63	3.48	3.37	3.29	3.23	3.18	3.14
10	4.10	3.71	3.48	3.33	3.22	3.14	3.07	3.02	2.98
12	3.89	3.49	3.26	3.11	3.00	2.91	2.85	2.80	2.75
15	3.68	3.29	3.06	2.90	2.79	2.71	2.64	2.59	2.54
16	3.63	3.24	3.01	2.85	2.74	2.66	2.59	2.54	2.49
18	3.55	3.16	2.93	2.77	2.66	2.58	2.51	2.46	2.41
20	3.49	3.10	2.87	2.71	2.60	2.51	2.45	2.39	2.35
21	3.47	3.07	2.84	2.68	2.57	2.49	2.42	2.37	2.32
24	3.40	3.01	2.78	2.62	2.51	2.42	2.36	2.30	2.25
25	3.39	2.99	2.76	2.60	2.49	2.40	2.34	2.28	2.24
27	3.35	2.96	2.73	2.57	2.46	2.37	2.31	2.25	2.20
28	3.34	2.95	2.71	2.56	2.45	2.36	2.29	2.24	2.19
30	3.32	2.92	2.69	2.53	2.42	2.33	2.27	2.21	2.16

	97.5 % percentile								
s	$r = 2$	3	4	5	6	7	8	9	10
2	39.00	39.17	39.25	39.30	39.33	39.36	39.37	39.39	39.40
3	16.04	15.44	15.10	14.88	14.73	14.62	14.54	14.47	14.42
4	10.65	9.98	9.60	9.36	9.20	9.07	8.98	8.90	8.84
5	8.43	7.76	7.39	7.15	6.98	6.85	6.76	6.68	6.62
6	7.26	6.60	6.23	5.99	5.82	5.70	5.60	5.52	5.46
7	6.54	5.89	5.52	5.29	5.12	4.99	4.90	4.82	4.76
8	6.06	5.42	5.05	4.82	4.65	4.53	4.43	4.36	4.30
9	5.71	5.08	4.72	4.48	4.32	4.20	4.10	4.03	3.96
10	5.46	4.83	4.47	4.24	4.07	3.95	3.85	3.78	3.72
12	5.10	4.47	4.12	3.89	3.73	3.61	3.51	3.44	3.37
15	4.77	4.15	3.80	3.58	3.41	3.29	3.20	3.12	3.06
16	4.69	4.08	3.73	3.50	3.34	3.22	3.12	3.05	2.99
18	4.56	3.95	3.61	3.38	3.22	3.10	3.01	2.93	2.87
20	4.46	3.86	3.51	3.29	3.13	3.01	2.91	2.84	2.77
21	4.42	3.82	3.48	3.25	3.09	2.97	2.87	2.80	2.73
24	4.32	3.72	3.38	3.15	2.99	2.87	2.78	2.70	2.64
25	4.29	3.69	3.35	3.13	2.97	2.85	2.75	2.68	2.61
27	4.24	3.65	3.31	3.08	2.92	2.80	2.71	2.63	2.57
28	4.22	3.63	3.29	3.06	2.90	2.78	2.69	2.61	2.55
30	4.18	3.59	3.25	3.03	2.87	2.75	2.65	2.57	2.51

Tabell 6: Critical values c for the Wilcoxon signed rank test, where n is the sample size and $C = n(n+1) - c$ [e.g., if $n = 20$, then $P(W \leq 61) = P(W \geq 149) \approx 0.05$]

n	0.025	0.05	$n(n+1)/2$	n	0.025	0.05	$n(n+1)/2$
5	0	1	15	18	41	48	171
6	1	3	21	19	47	54	190
7	3	4	28	20	53	61	210
8	4	6	36	21	59	68	231
9	6	9	45	22	67	76	253
10	9	11	55	23	74	84	276
11	11	14	66	24	82	92	300
12	14	18	78	25	90	101	325
13	18	22	91	26	99	111	351
14	22	26	105	27	108	120	378
15	26	31	120	28	117	131	406
16	30	36	136	29	127	141	435
17	35	42	153	30	138	152	465

Tabell 7: Critical values c for the Wilcoxon rank sum test, where m is the size of the smaller sample, and $C = m(m+n+1) - c$ [e.g., if $m = 4$ and $n = 8$, then $P(W \leq 16) = P(W \geq 36) \approx 0.05$]

n	$P(W \leq c)$	$m = 2$	3	4	5	6	7	8	9	10	11
2	0.025	3									
	0.05	3									
3	0.025	3	3								
	0.05	6	7								
4	0.025	3	6	11							
	0.05	3	7	12							
5	0.025	3	7	12	18						
	0.05	4	8	13	20						
6	0.025	3	8	13	19	27					
	0.05	4	9	14	21	29					
7	0.025	3	8	14	21	28	37				
	0.05	4	9	15	22	30	40				
8	0.025	4	9	15	22	30	39	50			
	0.05	5	10	16	24	32	42	52			
9	0.025	4	9	15	23	32	41	52	63		
	0.05	5	11	17	25	34	44	55	67		
10	0.025	4	10	16	24	33	43	54	66	79	
	0.05	5	11	18	27	36	46	57	70	83	
11	0.025	5	10	17	25	35	45	56	69	82	97
	0.05	5	12	19	28	38	48	60	73	87	101