

Försättsblad till skriftlig tentamen vid Linköpings universitet



Datum för tentamen	2016-04-01
Sal (1)	<u>KÅRA</u>
Tid	8-12
Kurskod	732G60
Provkod	TEN1
Kursnamn/benämning Provnamn/benämning	Statistiska metoder Tentamen
Institution	IDA
Antal uppgifter som ingår i tentamen	4
Jour/Kursansvarig Ange vem som besöker salen	Lotta Hallberg
Telefon under skrivtiden	013-281657
Besöker salen ca klockan	10
Kursadministratör/kontaktperson (namn + tfnr + mailaddress)	Annelie Almquist
Tillåtna hjälpmedel	Med skrivningen häftad formelsamling. Räknedosa.
Övrigt	
Antal exemplar i påsen	

TENTAMEN I STATISTISKA METODER 2016-04-01

Skrivtid: 8-12
Hjälpmedel: Med skrivningen häftad formelsamling. Räknedosa.
Jourhavande lärare: Lotta Hallberg

Redovisa och motivera kort alla dina lösningar

1

En livsmedelsaffär vill uppskatta hur mycket varje kund handlar för i snitt vid varje handlingstillfälle. Man noterar därför hur mycket 12 slumpmässigt valda kunder handlar för i kr. Dessutom noterar man kön på kunderna. Hur mycket en kund handlar för kan antas vara normalfördelat.

Kund nr	1	2	3	4	5	6	7	8	9	10	11	12
Antal kr	77	315	127	307	35	42	311	257	289	230	144	38
Kön	man	kvinn	kvinn	kvinn	kvinn	man	man	kvinn	man	kvinn	man	man

- a) Rita ett lådagram över hur mycket kunderna handlar för. Ange de olika värdena som används i lådagrammet. 2p
- b) Beräkna ett 95% konfidensintervall för hur mycket en kund handlar för i snitt. Tolka. 2p
- c) Någon hävdar att kvinnor handlar för mer än männen. Pröva på 5% signifikansnivå om så är fallet. Anta att standardavvikelseerna egentligen är lika för de båda grupperna. Tolka resultatet. 3p

2

För att undersöka om det är skillnad i rökvana och inkomstklass tillfrågades 356 personer i dessa frågor. Resultatet presenteras i följande korstabell med frekvenser:

Rökvana	Inkomstklass		
	Hög	Mellan	Låg
Är rökare	51	22	43
Har tidigare rökt	92	21	28
Har aldrig rökt	68	9	22

- a) Redovisa materialet i tre stapeldiagram bredvid varandra, ett för vardera inkomstklass och där staplarna för rökvana är angivna i procent. Är det någon skillnad i andelen rökare för de tre inkomstklasserna? Svara på frågan genom att tolka stapeldiagrammen. 2p
- b) Pröva med ett χ^2 -test på 5% signifikansnivå om det finns någon skillnad i rökvana för de tre inkomstklasserna. 3p

3

Kurt säljer påskägg på torget. Vikten hos ett ägg ska vara 200g i genomsnitt och standardavvikelsen är 10g. Vikten kan antas vara normalfördelad.

- a) Beräkna sannolikheten att ett slumpmässigt valt ägg har en vikt som ligger inom en standardavvikelse från 200g. Det vill säga sannolikheten att vikten ligger mellan 190 och 210g. 2p
- b) Kalle köper 5 ägg till sina 5 barn. Han väger dem och finner att medelvikten är 190g. Pröva hypotesen att medelvärdet är 200g till fördel för mothypotesen att medelvikten är mindre än 200g på 5% signifikansnivå genom att beräkna p-värdet. (Använd z, inte t) 2p

4

Följande data visar för 7 kvinnliga löpare deras genomsnittliga antal steg per sekund 'Steg' och springhastigheten i m/s 'm/s'. En bra löpare ökar antalet steg per sekund med hastigheten.

Steg	m/s
3.05	4.76
3.12	5.06
3.17	5.25
3.25	5.59
3.36	5.99
3.46	6.32
3.55	6.63

Regression Analysis: Steg versus m/s

Model Summary

S	R-sq	R-sq(adj)	R-sq(pred)
0,0092728	99,79%	99,75%	99,51%

Coefficients

Term	Coef	SE Coef	T-Value	P-Value	VIF
Constant	1,7653	0,0314			
m/s	0,26775	0,00551			

Regression Equation

Steg = 1,7653 + 0,26775 m/s

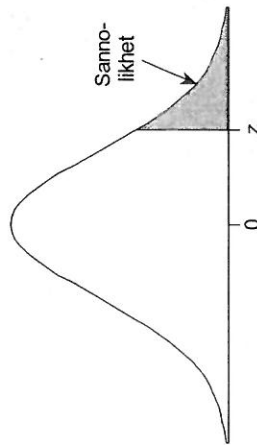
- a) Pröva om lutningskoefficienten ($b=0,26775$) är signifikant skild från 0 på 1% signifikansnivå. Tolka värdet på b. 2p
- b) Hur stor är förklaringsgraden= R^2 . Tolka R^2 ? 1p
- c) Hur många steg per sekund kan man förvänta sig att en löpare tar i snitt om hastigheten är 6.5m/s? 1p

Tabell 1 Normalfördelningen

Tabellen ger sannolikheten att en standardiserad normalfördelad variabel är större än ett visst värde z .

Exempel: Sannolikheten att få ett värde större än 1,23 är 10,9%.

Den standardiserade normalfördelningen är symmetrisk kring punkten noll. Därför ges sannolikheterna endast för positiva z -värden.

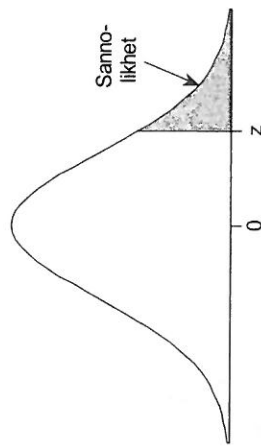


z	0	1	2	3	4	5	6	7	8	9
0,0	50,0	49,6	49,2	48,8	48,4	48,0	47,6	47,2	46,8	46,4
0,1	46,0	45,6	45,2	44,8	44,4	44,0	43,6	43,3	42,9	42,5
0,2	42,1	41,7	41,3	40,9	40,5	40,1	39,7	39,4	39,0	38,6
0,3	38,2	37,8	37,4	37,1	36,7	36,3	35,9	35,6	35,2	34,8
0,4	34,5	34,1	33,7	33,4	33,0	32,6	32,3	31,9	31,6	31,2
0,5	30,9	30,5	30,2	29,8	29,5	29,1	28,8	28,4	28,1	27,8
0,6	27,4	27,1	26,8	26,4	26,1	25,8	25,5	25,1	24,8	24,5
0,7	24,2	23,9	23,6	23,3	23,0	22,7	22,4	22,1	21,8	21,5
0,8	21,2	20,9	20,6	20,3	20,0	19,8	19,5	19,2	18,9	18,7
0,9	18,4	18,1	17,9	17,6	17,4	17,1	16,9	16,6	16,4	16,1
1,0	15,9	15,6	15,4	15,2	14,9	14,7	14,5	14,2	14,0	13,8
1,1	13,6	13,3	13,1	12,9	12,7	12,5	12,3	12,1	11,9	11,7
1,2	11,5	11,3	11,1	10,9	10,7	10,6	10,4	10,2	10,0	9,9
1,3	9,7	9,5	9,3	9,2	9,0	8,9	8,7	8,5	8,4	8,2
1,4	8,1	7,9	7,8	7,6	7,5	7,4	7,2	7,1	6,9	6,8
1,5	6,7	6,6	6,4	6,3	6,2	6,1	5,9	5,8	5,7	5,6
1,6	5,5	5,4	5,3	5,2	5,1	4,9	4,8	4,7	4,6	4,6
1,7	4,5	4,4	4,3	4,2	4,1	4,0	3,9	3,8	3,8	3,7
1,8	3,6	3,5	3,4	3,4	3,3	3,2	3,1	3,1	3,0	2,9
1,9	2,9	2,8	2,7	2,7	2,6	2,6	2,5	2,4	2,4	2,3
2,0	2,3	2,2	2,2	2,1	2,1	2,0	2,0	1,9	1,9	1,8
2,1	1,8	1,7	1,7	1,7	1,6	1,6	1,5	1,5	1,5	1,4
2,2	1,4	1,4	1,3	1,3	1,3	1,2	1,2	1,2	1,1	1,1
2,3	1,1	1,0	1,0	1,0	1,0	0,9	0,9	0,9	0,9	0,8
2,4	0,8	0,8	0,8	0,8	0,7	0,7	0,7	0,7	0,7	0,6

z	0	1	2	3	4	5	6	7	8	9
2,5	0,6	0,6	0,6	0,6	0,6	0,5	0,5	0,5	0,5	0,5
2,6	0,5	0,5	0,4	0,4	0,4	0,4	0,4	0,4	0,4	0,4
2,7	0,3	0,3	0,3	0,3	0,3	0,3	0,3	0,3	0,3	0,3
2,8	0,3	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,2
2,9	0,2	0,2	0,2	0,2	0,2	0,2	0,2	0,1	0,1	0,1
3,0	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1	0,1

Tabell 2 Normalfördelningen

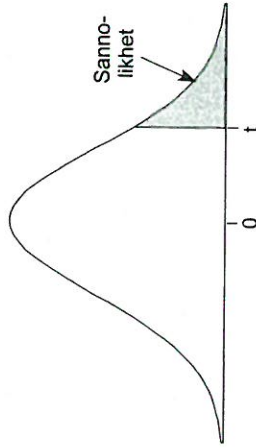
z -värden för vissa sannolikheter.



Sannolikhet	z	Sannolikhet	z	Sannolikhet	z
50%	0,00	9%	1,34	2,5%	1,96
40%	0,25	8%	1,41	1,0%	2,33
30%	0,52	7%	1,48	0,5%	2,58
20%	0,84	6%	1,55	0,1%	3,09
10%	1,28	5%	1,64	0,05%	3,29

Tabell 3 t-fördelningen

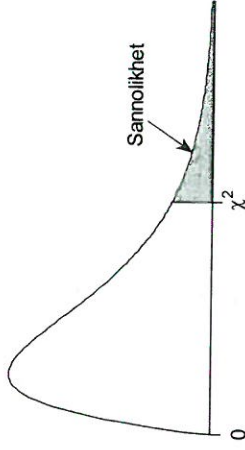
t-värden för vissa sannolikheter.



fg	Sannolikhet					
	5%	2,5%	1%	0,5%	0,1%	0,05%
5	2,02	2,57	3,36	4,03	5,89	6,87
6	1,94	2,45	3,14	3,71	5,21	5,96
7	1,89	2,36	3,00	3,50	4,79	5,41
8	1,86	2,31	2,90	3,36	4,50	5,04
9	1,83	2,26	2,82	3,25	4,30	4,78
10	1,81	2,23	2,76	3,17	4,14	4,59
11	1,80	2,20	2,72	3,11	4,02	4,44
12	1,78	2,18	2,68	3,05	3,93	4,32
13	1,77	2,16	2,65	3,01	3,85	4,22
14	1,76	2,14	2,62	2,98	3,79	4,14
15	1,75	2,13	2,60	2,95	3,73	4,07
16	1,75	2,12	2,58	2,92	3,69	4,01
17	1,74	2,11	2,57	2,90	3,65	3,97
18	1,73	2,10	2,55	2,88	3,61	3,92
19	1,73	2,09	2,54	2,86	3,58	3,88
20	1,72	2,09	2,53	2,85	3,55	3,85
21	1,72	2,08	2,52	2,83	3,53	3,82
22	1,72	2,07	2,51	2,82	3,50	3,79
23	1,71	2,07	2,50	2,81	3,48	3,77
24	1,71	2,06	2,49	2,80	3,47	3,75
25	1,71	2,06	2,49	2,79	3,45	3,73
26	1,71	2,06	2,48	2,78	3,44	3,71
27	1,70	2,05	2,47	2,77	3,42	3,69
28	1,70	2,05	2,47	2,76	3,41	3,67
29	1,70	2,05	2,46	2,76	3,40	3,66
30	1,70	2,04	2,46	2,75	3,39	3,65
40	1,68	2,02	2,42	2,70	3,31	3,55
60	1,67	2,00	2,39	2,66	3,23	3,46
120	1,66	1,98	2,36	2,62	3,16	3,37
∞	1,64	1,96	2,33	2,58	3,09	3,29

Tabell 4 χ^2 -fördelningen

χ^2 -värden för vissa sannolikheter.



fg	Sannolikhet			fg	Sannolikhet		
	5%	1%	0,1%		5%	1%	0,1%
1	3,84	6,63	10,83	26	38,89	45,64	54,05
2	5,99	9,21	13,82	27	40,11	46,96	55,48
3	7,81	11,34	16,27	28	41,34	48,28	56,89
4	9,49	13,28	18,47	29	42,56	49,59	58,30
5	11,07	15,09	20,52	30	43,77	50,89	59,70
6	12,59	16,81	22,46	31	44,99	52,19	61,10
7	14,07	18,48	24,32	32	46,19	53,49	62,49
8	15,51	20,09	26,12	33	47,40	54,78	63,87
9	16,92	21,67	27,88	34	48,60	56,06	65,25
10	18,31	23,21	29,59	35	49,80	57,34	66,62
11	19,68	24,72	31,26	36	51,00	58,62	67,99
12	21,03	26,22	32,91	37	52,19	59,89	69,35
13	22,36	27,69	34,53	38	53,38	61,16	70,71
14	23,68	29,14	36,12	39	54,57	62,43	72,06
15	25,00	30,58	37,70	40	55,76	63,69	73,41
16	26,30	32,00	39,25	41	56,94	64,95	74,75
17	27,59	33,41	40,79	42	58,12	66,21	76,09
18	28,87	34,81	42,31	43	59,30	67,46	77,41
19	30,14	36,19	43,82	44	60,48	68,71	78,75
20	31,41	37,57	45,31	45	61,66	69,96	80,08
21	32,67	38,93	46,80	46	62,83	71,20	81,39
22	33,92	40,29	48,27	47	64,00	72,44	82,72
23	35,17	41,64	49,73	48	65,17	73,68	84,03
24	36,42	42,98	51,18	49	66,34	74,92	85,35
25	37,65	44,31	52,62	50	67,50	76,15	86,66

Beskrivande statistik

Medelvärde

$$\bar{x} = \frac{\sum f \cdot x}{n}$$

Standardavvikelsen

$$s = \sqrt{\frac{\sum x^2 - \frac{(\sum x)^2}{n}}{n-1}}$$

Regressionslinjen

$$y = a + b \cdot x$$

$$a = \bar{y} - b \cdot \bar{x}$$
$$b = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sum x^2 - \frac{(\sum x)^2}{n}}$$

Korrelationskoefficienten

$$r = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\sqrt{\left(\sum x^2 - \frac{(\sum x)^2}{n}\right) \left(\sum y^2 - \frac{(\sum y)^2}{n}\right)}}$$

Residualspridningen

$$s_e = \sqrt{\frac{(n-1) \cdot s_y^2 \cdot (1-r^2)}{n-2}}$$

Konfidsintervall

Ett stickprov, populationens andel

$$p \pm z \cdot \sqrt{\frac{p(1-p)}{n}}$$

Ett stickprov, populationens medelvärde

$$\bar{x} \pm z \cdot \frac{s}{\sqrt{n}} \quad (\text{stort stickprov})$$

$$\bar{x} \pm t \cdot \frac{s}{\sqrt{n}} \quad (\text{normalfördelad variabel, } n - 1 \text{ frihetsgrader})$$

Två oberoende stickprov, skillnader mellan andelar

$$(p_1 - p_2) \pm z \cdot \sqrt{\frac{p_1(1-p_1)}{n_1} + \frac{p_2(1-p_2)}{n_2}}$$

Två oberoende stickprov, skillnader mellan medelvärden

$$(\bar{x}_1 - \bar{x}_2) \pm z \cdot \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \quad (\text{stora oberoende stickprov})$$

$$(\bar{x}_1 - \bar{x}_2) \pm t \cdot \sqrt{s_p^2 \left(\frac{1}{n_1} + \frac{1}{n_2} \right)} \quad \text{där}$$

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \quad (n_1 + n_2 - 2 \text{ frihetsgrader})$$

(oberoende stickprov, normalfördelade variabler med samma standardavvikelse i populationerna)

Hypotesprövning, testfunktioner

Ett stickprov, populationens andel

$$z = \frac{p - \pi_0}{\sqrt{\frac{\pi_0(1 - \pi_0)}{n}}}$$

Ett stickprov, skillnad mellan två andelar

$$z = \frac{p_1 - p_2}{\sqrt{\frac{p_1 + p_2}{n}}}$$

Ett stickprov, populationens medelvärde

$$z = \frac{\bar{x} - \mu_0}{s / \sqrt{n}} \quad t = \frac{\bar{x} - \mu_0}{s / \sqrt{n}}$$

Två oberoende stickprov, skillnader mellan andelar

$$z = \frac{p_1 - p_2}{\sqrt{p(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \quad (p \text{ är andelen i det sammanslagna stickprovet})$$

Två oberoende stickprov, skillnader mellan medelvärden

$$z = \frac{(\bar{x}_1 - \bar{x}_2) - d_0}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \quad t = \frac{(\bar{x}_1 - \bar{x}_2) - d_0}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

χ^2 -metoden

$$\chi^2 = \sum \frac{(O - E)^2}{E}$$

Tabeller & formler

SVANTE KÖRNER
LARS WAHLGREN

Supplement till:
Statistiska metoder
Tredje upplagan. Art.nr 6809