Stockholms universitet
Statistiska institutionen
Raul Cano

SKRIVNINGSDATUM: 2013-01-15

Skriftlig tentamen i **Statistikens grunder 2** (6 hp), ingående som moment 3 i kursen **Statistikens grunder, GN, 15 hp.**

Skrivtid: 5 timmar
Hjälpmedel: Miniräknare, språklexikon. Vidhäftade formel- och tabellblad.
Tentamensgenomgång och återlämning: Måndagen den 28 januari, kl. 18.00 i B319.
Därefter kan skrivningarna hämtas på studentexpeditionen, plan 7 i B-huset.

Tentamen består av fem uppgifter som kan ge totalt 100 poäng. För betyget A gäller 90-100 p., för betyget B gäller 80-89 p., för betyget C gäller 70-79 p., för betyget D gäller 60-69 p., för betyget E gäller 50-59 p., för betyget Fx gäller 40-49 p. och för betyget F gäller 0-39 p. För detaljerade betygskriterier se kursbeskrivningen på kurshemsidan.
För full poäng på en uppgift krävs fullständiga och väl motiverade lösningar.

Uppgift 1: (20 poäng)

Man har vägt fyra slumpmässigt utvalda personer. Resultat (vikt i kg) : 78, 80, 82, 80.
Antag att de uppmätta vikterna kan uppfattas som oberoende observationer på en normalfördelad stokastisk variabel med okänt väntevärde μ och okänd varians σ^2.

a) Beräkna ett 95%-igt konfidensintervall för μ och tolka resultatet i ord. (10 p.)

b) Testa att $\mu > 80$. Sätt upp hypoteser, ange testvariabel och beslutsregel. Använd signifikansnivån 5%. Vilken blir din slutsats? (10 p.)

Uppgift 2: (20 poäng)

I den stora staden New York har 400 av 500 tillfrågade (slumpmässigt utvalda) familjer svarat att de har köpt produkt X

a) Är andelen familjer (som har köpt produkt X) i New York större än 0,75 ? Sätt upp hypoteser, ange testvariabel och beslutsregel. Använd signifikansnivån 5%. Vilken blir din slutsats? (15 p.)

b) Beräkna p-värdet och förklara hur det kan användas för att genomföra hypotestestet. (5 p.)

Uppgift 3: (20 poäng)

För att mäta effekten av behandling A på blodsockernivån i jämförelse med behandling B väljer man slumpmässigt 10 personer med förhöjd blodsockernivå. Därefter väljs slumpmässigt bland de 10 personerna fem till behandling A, medan de övriga 5 erhåller behandling B. Finns det någon skillnad i effekt på blodsockret mellan de båda behandlingarna? Använd signifikansnivån 10 %.
Nedan ges mätvärdena för de patienterna som ingick i försöket.

Blodsockernivå					
Behandlad med A	10,9	8,1	8,5	8,0	9,6
Behandlad med B	9,7	5,7	8,2	6,9	8,4

Antag att blodsockernivåerna kan uppfattas som oberoende observationer på två normalfördelade oberoende stokastiska variabler med okända väntevärden μ_A och μ_B. Antag vidare att de okända varianserna σ_A^2 och σ_B^2 är lika stora. (20 p.)

Uppgift 4: (20 poäng)

Ägaren till ett varuhus påstår att 60% av kunderna inte köper något, utan bara tittar på varorna. Han påstår vidare att 30% köper en enda artikel och att 10 % köper minst två artiklar. För att undersöka om detta påstående är riktigt har en anställd observerat 100 slumpmässigt utvalda kunder med följande resultat:

Antal köpta artiklar	Antal kunder
0	80
1	10
Minst 2	10

Är ägarens påstående riktigt? (använd 10 % signifikansnivå).

Uppgift 5: (20 poäng)

I ett företag arbetar man med ett projekt som syftar till att utveckla 4 nya produkter och (efter genomförande av en statistisk analys) välja en av de för fastproduktion. Företagsledningen har därför fyra alternativ att välja mellan: Produkt A, B, C och D. Lönsamheten för de olika produkterna är beroende av hur efterfrågan på de ingående produkterna blir. Ledningen har betraktat endast 4 olika typer av efterfrågan för varje produkt: E1, E2, E3 och E4. Marknadsundersökningsenheten (inom företaget) har bekräftat att om man väljer Produkt A räknar man med en vinst om 10 miljoner kronor (Mk) oavsett vilken typ av efterfrågan inträffar, medan om man väljer Produkt B blir vinsten 18 Mk, 18 Mk, 2 Mk respektive 10 Mk för efterfrågan E1, E2, E3 respektive E4. Motsvarande vinst för Produkt C är 2 Mk, 30 Mk, 0 Mk och 0 Mk och för Produkt D 10 Mk, 25 Mk, 2 Mk och 2 Mk.

a) Bestäm med hjälp av Laplacekriteriet vilken produkt Företagsledningen bör välja. (10 p.)

b) Antag att sannolikheterna för att de olika typer av efterfrågan inträffar är 0.2, 0.4, 0.3 respektive 0.1 för E1, E2, E3 respektive E4 och bestäm vilken produkt Företagsledningen bör välja. (10 p.)

Formler

Räkneregler för väntevärden och varianser där X och Y är stokastiska variabler, a, b och c är konstanter:

$$
\begin{aligned}
E(c) &= c \\
E(cX) &= cE(X) \\
E(c+X) &= c+E(X) \\
E(aX+bY) &= aE(X)+bE(Y) \\
V(c) &= 0 \\
V(cX) &= c^2V(X) \\
V(c+X) &= V(X) \\
V(aX+bY) &= a^2V(X)+b^2V(Y)+2abCov\,(X,Y)
\end{aligned}
$$

Väntevärde och varians för urvalsmedelvärdet $\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n}$ där alla $X_1, X_2, ..., X_n$ är oberoende och har väntevärde μ och varians σ^2:

$$
\begin{aligned}
E\left(\overline{X}\right) &= \mu \\
V\left(\overline{X}\right) &= \frac{\sigma^2}{n}
\end{aligned}
$$

Ändlighetskorrektion:

$$\frac{N-n}{N-1}$$

Testvariabler:

$$Z = \frac{\overline{X}-\mu_0}{\sigma/\sqrt{n}}$$

$$t = \frac{\overline{X}-\mu_0}{S/\sqrt{n}}$$

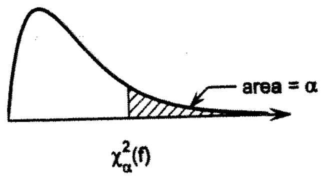
$$Z = \frac{\overline{X}-\overline{Y}-D_0}{\sqrt{\frac{\sigma_1^2}{n}+\frac{\sigma_2^2}{m}}}$$

$$t = \frac{\overline{X}-\overline{Y}-D_0}{\sqrt{S_p^2\left(\frac{1}{n}+\frac{1}{m}\right)}}, \quad \text{där } S_p^2 = \frac{(n-1)\,S_1^2+(m-1)\,S_2^2}{n+m-2}$$

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

Tabell 4. χ^2**-fördelningen**

$P(X > \chi^2_\alpha(f)) = \alpha$ där $X \in \chi^2(f)$



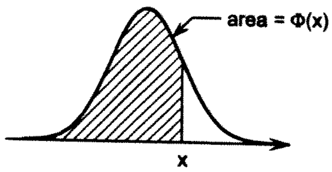
f	α 0.9995	0.999	0.995	0.99	0.975	0.95	0.05	0.025	0.01	0.005	0.001	0.0005
1	0.00	0.00	0.00	0.00	0.00	0.00	3.84	5.02	6.63	7.88	10.83	12.12
2	0.00	0.00	0.01	0.02	0.05	0.10	5.99	7.38	9.21	10.60	13.82	15.20
3	0.02	0.02	0.07	0.11	0.22	0.35	7.81	9.35	11.34	12.84	16.27	17.73
4	0.06	0.09	0.21	0.30	0.48	0.71	9.49	11.14	13.28	14.86	18.47	20.00
5	0.16	0.21	0.41	0.55	0.83	1.15	11.07	12.83	15.09	16.75	20.52	22.11
6	0.30	0.38	0.68	0.87	1.24	1.64	12.59	14.45	16.81	18.55	22.46	24.10
7	0.48	0.60	0.99	1.24	1.69	2.17	14.07	16.01	18.48	20.28	24.32	26.02
8	0.71	0.86	1.34	1.65	2.18	2.73	15.51	17.53	20.09	21.95	26.12	27.87
9	0.97	1.15	1.73	2.09	2.70	3.33	16.92	19.02	21.67	23.59	27.88	29.67
10	1.26	1.48	2.16	2.56	3.25	3.94	18.31	20.48	23.21	25.19	29.59	31.42
11	1.59	1.83	2.60	3.05	3.82	4.57	19.68	21.92	24.72	26.76	31.26	33.14
12	1.93	2.21	3.07	3.57	4.40	5.23	21.03	23.34	26.22	28.30	32.91	34.82
13	2.31	2.62	3.57	4.11	5.01	5.89	22.36	24.74	27.69	29.82	34.53	36.48
14	2.70	3.04	4.07	4.66	5.63	6.57	23.68	26.12	29.14	31.32	36.12	38.11
15	3.11	3.48	4.60	5.23	6.26	7.26	25.00	27.49	30.58	32.80	37.70	39.72
16	3.54	3.94	5.14	5.81	6.91	7.96	26.30	28.85	32.00	34.27	39.25	41.31
17	3.98	4.42	5.70	6.41	7.56	8.67	27.59	30.19	33.41	35.72	40.79	42.88
18	4.44	4.90	6.26	7.01	8.23	9.39	28.87	31.53	34.81	37.16	42.31	44.43
19	4.91	5.41	6.84	7.63	8.91	10.12	30.14	32.85	36.19	38.58	43.82	45.97
20	5.40	5.92	7.43	8.26	9.59	10.85	31.41	34.17	37.57	40.00	45.31	47.50
21	5.90	6.45	8.03	8.90	10.28	11.59	32.67	35.48	38.93	41.40	46.80	49.01
22	6.40	6.98	8.64	9.54	10.98	12.34	33.92	36.78	40.29	42.80	48.27	50.51
23	6.92	7.53	9.26	10.20	11.69	13.09	35.17	38.08	41.64	44.18	49.73	52.00
24	7.45	8.08	9.89	10.86	12.40	13.85	36.42	39.36	42.98	45.56	51.18	53.48
25	7.99	8.65	10.52	11.52	13.12	14.61	37.65	40.65	44.31	46.93	52.62	54.95
26	8.54	9.22	11.16	12.20	13.84	15.38	38.89	41.92	45.64	48.29	54.05	56.41
27	9.09	9.80	11.81	12.88	14.57	16.15	40.11	43.19	46.96	49.64	55.48	57.86
28	9.66	10.39	12.46	13.56	15.31	16.93	41.34	44.46	48.28	50.99	56.89	59.30
29	10.23	10.99	13.12	14.26	16.05	17.71	42.56	45.72	49.59	52.34	58.30	60.73
30	10.80	11.59	13.79	14.95	16.79	18.49	43.77	46.98	50.89	53.67	59.70	62.16
40	16.91	17.92	20.71	22.16	24.43	26.51	55.76	59.34	63.69	66.77	73.40	76.09
50	23.46	24.67	27.99	29.71	32.36	34.76	67.50	71.42	76.15	79.49	86.66	89.56
60	30.34	31.74	35.53	37.48	40.48	43.19	79.08	83.30	88.38	91.95	99.61	102.69
70	37.47	39.04	43.28	45.44	48.76	51.74	90.53	95.02	100.43	104.21	112.32	115.58
80	44.79	46.52	51.17	53.54	57.15	60.39	101.88	106.63	112.33	116.32	124.84	128.26
90	52.28	54.16	59.20	61.75	65.65	69.13	113.15	118.14	124.12	128.30	137.21	140.78
100	59.90	61.92	67.33	70.06	74.22	77.93	124.34	129.56	135.81	140.17	149.45	153.17

Tabeller

Tabell 1. Standardiserad normalfördelning

$\Phi(x) = P(X \leq x)$ där $X \in N(0, 1)$

För negativa värden, utnyttja att $\Phi(x) = 1 - \Phi(-x)$

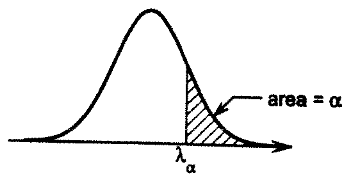


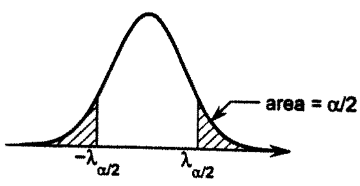
x	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.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	.97725	.97778	.97831	.97882	.97932	.97982	.98030	.98077	.98124	.98169
2.1	.98214	.98257	.98300	.98341	.98382	.98422	.98461	.98500	.98537	.98574
2.2	.98610	.98645	.98679	.98713	.98745	.98778	.98809	.98840	.98870	.98899
2.3	.98928	.98956	.98983	.99010	.99036	.99061	.99086	.99111	.99134	.99158
2.4	.99180	.99202	.99224	.99245	.99266	.99286	.99305	.99324	.99343	.99361
2.5	.99379	.99396	.99413	.99430	.99446	.99461	.99477	.99492	.99506	.99520
2.6	.99534	.99547	.99560	.99573	.99585	.99598	.99609	.99621	.99632	.99643
2.7	.99653	.99664	.99674	.99683	.99693	.99702	.99711	.99720	.99728	.99736
2.8	.99744	.99752	.99760	.99767	.99774	.99781	.99788	.99795	.99801	.99807
2.9	.99813	.99819	.99825	.99831	.99836	.99841	.99846	.99851	.99856	.99861
3.0	.99865									
3.1	.99903									
3.2	.99931									
3.3	.99952									
3.4	.99966									
3.5	.99977									
3.6	.99984									
3.7	.99989									
3.8	.99993									
3.9	.99995									
4.0	.99997									

Tabell 2. Normalfördelningens kvantiler

$P(X > \lambda_\alpha) = \alpha$ där $X \in N(0, 1)$

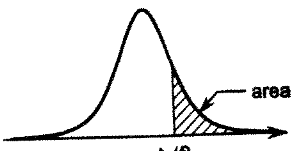
α	λ_α	α	λ_α
0.1	1.2816	0.001	3.0902
0.05	1.6449	0.0005	3.2905
0.025	1.9600	0.0001	3.7190
0.01	2.3263	0.00005	3.8906
0.005	2.5758	0.00001	4.2649





Tabell 3. *t***-fördelningen**

$P(X > t_\alpha(f)) = \alpha$ där $X \in t(f)$



f \ α	0.1	0.05	0.025	0.01	0.005	0.001	0.0005
1	3.08	6.31	12.71	31.82	63.66	318.31	636.62
2	1.89	2.92	4.30	6.96	9.92	22.33	31.60
3	1.64	2.35	3.18	4.54	5.84	10.21	12.92
4	1.53	2.13	2.78	3.75	4.60	7.17	8.61
5	1.48	2.02	2.57	3.36	4.03	5.89	6.87
6	1.44	1.94	2.45	3.14	3.71	5.21	5.96
7	1.41	1.89	2.36	3.00	3.50	4.79	5.41
8	1.40	1.86	2.31	2.90	3.36	4.50	5.04
9	1.38	1.83	2.26	2.82	3.25	4.30	4.78
10	1.37	1.81	2.23	2.76	3.17	4.14	4.59
11	1.36	1.80	2.20	2.72	3.11	4.02	4.44
12	1.36	1.78	2.18	2.68	3.05	3.93	4.32
13	1.35	1.77	2.16	2.65	3.01	3.85	4.22
14	1.35	1.76	2.14	2.62	2.98	3.79	4.14
15	1.34	1.75	2.13	2.60	2.95	3.73	4.07
16	1.34	1.75	2.12	2.58	2.92	3.69	4.01
17	1.33	1.74	2.11	2.57	2.90	3.65	3.97
18	1.33	1.73	2.10	2.55	2.88	3.61	3.92
19	1.33	1.73	2.09	2.54	2.86	3.58	3.88
20	1.33	1.72	2.09	2.53	2.85	3.55	3.85
21	1.32	1.72	2.08	2.52	2.83	3.53	3.82
22	1.32	1.72	2.07	2.51	2.82	3.50	3.79
23	1.32	1.71	2.07	2.50	2.81	3.48	3.77
24	1.32	1.71	2.06	2.49	2.80	3.47	3.75
25	1.32	1.71	2.06	2.49	2.79	3.45	3.73
26	1.31	1.71	2.06	2.48	2.78	3.43	3.71
27	1.31	1.70	2.05	2.47	2.77	3.42	3.69
28	1.31	1.70	2.05	2.47	2.76	3.41	3.67
29	1.31	1.70	2.05	2.46	2.76	3.40	3.66
30	1.31	1.70	2.04	2.46	2.75	3.39	3.65
40	1.30	1.68	2.02	2.42	2.70	3.31	3.55
60	1.30	1.67	2.00	2.39	2.66	3.23	3.46
120	1.29	1.66	1.98	2.36	2.62	3.16	3.37
∞	1.28	1.64	1.96	2.33	2.58	3.09	3.29

32/32

Statistiska institutionen

Stockholms universitet

Rättningsblad

Datum: 15/1 - 2013

Sal: Brunnsvikssalen

Tenta: Statistikens grunder 2

Kurs: Statistikens grunder, kväll

ANONYMKOD: [redacted]

OBS! SKRIV ÄVEN PÅ BAKSIDAN AV SKRIVBLADEN

Markera besvarade uppgifter med kryss

1	2	3	4	5	6	7	8	9	Antal inl. blad
X	X	X	X	X					3
Lär.ant. 20p	18p	20p	20p	20p					

SE

POÄNG	BETYG	Lärarens sign.
98	A	RC

SU, STATISTIK

Skrivsal: Brunnvikssalen Anonymkod [redacted] Blad nr: 1

1) 20p

~~Uppgift 1~~

a) $\bar{x} = \frac{78+80+82+80}{4} = 80$ R $fg: n-1 = 4-1 = 3$

$\alpha = 0{,}05 \quad \alpha/2 = 0{,}025$

$s^2 = \frac{1}{4-1}\left((78-80)^2 + (80-80)^2 + (82-80)^2 + (80-80)^2\right) =$

$= \frac{1}{3}(4+4) = \frac{1}{3}(8) \approx 2{,}6667$ R

$s \approx \sqrt{2{,}6667} \approx 1{,}6330$ R

Eftersom n är litet (< 30), men jag också vet att variabeln är normalfördelad kan jag använda en t-fördelning i konstruktionen av k.i. R

k.i. $\mu = \bar{x} \pm t_{\alpha/2}^{(n-1)fg} \cdot \frac{s}{\sqrt{n}} = 80 \pm 3{,}18 \cdot \frac{1{,}633}{\sqrt{4}} = 80 \pm 2{,}59647$ R

$\mu_u = 80 - 2{,}59647 = 77{,}40353 \approx 77{,}40$

$\mu_ö = 80 + 2{,}59647 = 82{,}59647 \approx 82{,}60$

k.i. μ: [77,40 ; 82,60]

b) $H_0: \mu \leq 80 \quad H_1: \mu > 80 \qquad \alpha = 0{,}05 \qquad n-1\ fg$ R

$t_{OBS} = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{80-80}{1{,}6330/\sqrt{4}} = 0$

Förkasta H_0 om $t_{OBS} > t_{\alpha}^{(n-1)fg} = t_{0,05}^{(3)}$

från tabell: $t_{0,05}^{(3)} = 2{,}35$, R alltså $t_{OBS} \not> t_{0,05}^{(3)}$ och vi kan inte förkasta H_0 på 5%-nivån.

R

Uppgift 2.

a.)

$\hat{p} = \frac{400}{500} = 0,8$ ∂) 18 ℓ $\alpha = 0,05$

$H_0 : \pi = \pi \leq 0,75$ $H_1 : \pi > 0,75$

Stort n (>30) gör att vi kan approximera den binominalfördelning vi utgår ifrån, med en normalfördelning.

$$Z_{OBS} = \frac{\hat{p} - \pi_0}{\sqrt{\frac{\pi_0(1-\pi_0)}{n}}} = \frac{0,8 - 0,75}{\sqrt{\frac{0,75 \cdot 0,25}{500}}} \approx 2,58$$

F (−2p)

Förkasta H_0 om $Z_{OBS} > Z_{0,05}$.

från tabell: $Z_{0,05} = 1,6449$ R Alltså gäller $Z_{OBS} = 2,58 > 1,6449 = Z_{0,05}$ och vi bör förkasta H_0 på 5%-nivån. ~~Vi har fått anledning att tro att $\pi > 0,75$~~

b.) p-värdet anger den lägsta nivå då vi fortfarande måste förkasta H_0.

$p = P(Z_{OBS} > Z \mid H_0 \text{ sann}) \approx 1 - \Phi(2,58) = 1 - 0,99506$

$= 0,00495$

Vi kunde genomföra hypotestestet genom att observera att: $0,00495 < 0,05$.

OK

Förkastas H_0 redan på 0,495%-nivån, gör den också det för den högre 5%-nivån

SU, STATISTIK

Skrivsal: Brunnvikssalen Anonymkod: [redacted] Blad nr: 2

3) 20p

Uppgift 3. $\alpha = 0,1 \quad \alpha/2 = 0,05$

$$\bar{x}_A = \frac{10,9+8,1+8,5+8+9,6}{5} = 9,02$$

$$\bar{x}_B = \frac{9,7+5,7+8,2+6,9+8,4}{5} = 7,78$$

$$s_A^2 = \frac{1}{5-1}\left((10,9-9,02)^2+(8,1-9,02)^2+(8,5-9,02)^2+(8-9,02)^2+(9,6-9,02)^2\right) =$$

$$= \frac{1}{4}\left(3,5344+0,8464+0,2704+1,0404+0,3364\right) =$$

$$= \frac{1}{4}(6,028) = 1,507$$

$$s_B^2 = \frac{1}{5-1}\left((9,7-7,78)^2+(5,7-7,78)^2+(8,2-7,78)^2+(6,9-7,78)^2+(8,4-7,78)^2\right)$$

$$= \frac{1}{4}\left(3,6864+4,3264+0,42+0,7744+0,3844\right) =$$

$$= \frac{1}{4}(9,5916) = 2,3979$$

$$s_p^2 = \frac{(5-1)1,507+(5-1)2,3979}{5+5-2} = \frac{15,6196}{8} = 1,95245$$

$H_0: \mu_A = \mu_B \qquad H_1: \mu_A \neq \mu_B$

$$t_{OBS} = \frac{\bar{x}_A - \bar{x}_B - D_0}{\sqrt{s_p^2\left(\frac{1}{n}+\frac{1}{m}\right)}} = \frac{9,02-7,78-0}{\sqrt{1,95245\left(\frac{1}{5}+\frac{1}{5}\right)}} \approx 1,40$$

Förkasta H_0 om $t_{OBS} > t_{\alpha/2}^{(n+m-2)} = t_{0,05}^{(8)}$ eller $t_{OBS} < -t_{0,05}^{(8)}$

Från tabell: $t_{0,05}^{(8)} = 1,86$ så att $-t_{0,05}^{(8)} = -1,86$

Men $t_{OBS} \not> t_{0,05}^{(8)}$ och $t_{OBS} \not< -t_{0,05}^{(8)}$

Vi kan därför inte förkasta H_0 på 10%-nivån.

Vår observation är inte tillräckligt osannolik då H_0 är sann för att kunna förkasta H_0 på 10%-nivån. De olika behandlingarna har inte visat sig ha olika effekt enligt testet vi genomfört.

Uppgift 4. $\alpha = 0{,}1$ $\alpha/2 = 0{,}05$

4) 2 p

Antal köpta artiklar	Antal kunder
0	80 (60)
1	10 (30)
Minst 2	10 (10)
Totalt:	100

Inom parentesen har jag skrivit det förväntade antalet kunder enligt ägarens påstående.

Dvs: De som bara tittar: 60% av 100 = 60

De som köper 1 vara: 30% av 100 = 30

De som köper minst 2: 10% av 100 = 10.

~~Jag undersöker sedan om ägarens påstående är förenligt~~ med den anställdes observation (på 10%-nivån).

$$\chi^2_{OBS} = \sum \frac{(O-E)^2}{E} = \frac{(80-60)^2}{60} + \frac{(10-30)^2}{30} + \frac{(10-10)^2}{10} =$$

$= 20$

$fg: (K-1)(R-1) = (2-1)(3-1) = 2$

Förkasta ägarens påstående om $\chi^2_{OBS} > \chi^{2(2)}_{0,05}$

Från tabell: $\chi^{2(2)}_{0,05} = 5{,}99$. Men $\chi^2_{OBS} = 20 > 5{,}99 = \chi^{2(2)}_{0,05}$

Vi bör därför förkasta ägarens påstående på 10%-nivån.

SU, STATISTIK

Skrivsal: Brunnsvikssalen   Anonymkod: [redacted]   Blad nr: 3

5) 20p

Uppgift 5

Laplacekriteriet uppmanar oss att sätta varje utfall till en lika stor sannolikhet.

I vårat fall: $P(E_1) = P(E_2) = P(E_3) = P(E_4) = \frac{1}{4}$

a) $E(A) = 10 \cdot \frac{1}{4} + 10 \cdot \frac{1}{4} + 10 \cdot \frac{1}{4} + 10 \cdot \frac{1}{4} = 10$

$E(B) = 18 \cdot \frac{1}{4} + 18 \cdot \frac{1}{4} + 2 \cdot \frac{1}{4} + 10 \cdot \frac{1}{4} = 12$

$E(C) = 2 \cdot \frac{1}{4} + 30 \cdot \frac{1}{4} + 0 \cdot \frac{1}{4} + 0 \cdot \frac{1}{4} = 8$

$E(D) = 10 \cdot \frac{1}{4} + 25 \cdot \frac{1}{4} + 2 \cdot \frac{1}{4} + 2 \cdot \frac{1}{4} = 9{,}75$

Sedan väljer vi alternativet med maximalt väntevärde, dvs B.

b) $E(A) = 10 \cdot 0{,}2 + 10 \cdot 0{,}4 + 10 \cdot 0{,}3 + 10 \cdot 0{,}1 = 10$

$E(B) = 18 \cdot 0{,}2 + 18 \cdot 0{,}4 + 2 \cdot 0{,}3 + 10 \cdot 0{,}1 = 12{,}4$

$E(C) = 2 \cdot 0{,}2 + 30 \cdot 0{,}4 + 0 \cdot 0{,}3 + 0 \cdot 0{,}1 = 12{,}4$

$E(D) = 10 \cdot 0{,}2 + 25 \cdot 0{,}4 + 2 \cdot 0{,}3 + 2 \cdot 0{,}1 = 12{,}8$

Maximalt väntevärde får vi vid alternativ D, det ger oss ett skäl att välja D. Men det finns andra rimliga kriterier att ta hänsyn till:

	E_1	E_2	E_3	E_4	min	max	max-regret
A	10	10	10	10	10	10	10-10=0
B	18	18	2	10	2	18	18-2=16
C	2	30	0	0	0	30	30-0=30
D	10	25	2	2	2	25	25-2=23

maximal utdelning kan vi få vid handling C, när E_2 inträffar (med 40% sannolikhet). en vågad (girig?) person kunde därför välja C. Enligt max-min-kriteriet ska vi välja A. Enligt min-maxregret-kriteriet bör vi välja A. Men dessa resonemang måste förfinas genom att ta hänsyn till de olika sannolikheterna [illegible]

(29)

Statistiska institutionen

Stockholms universitet

Rättningsblad

Datum: 15/1 - 2013

Sal: Brunnsvikssalen

Tenta: Statistikens grunder 2

Kurs: Statistikens grunder, kväll

ANONYMKOD: [redacted]

OBS! SKRIV ÄVEN PÅ BAKSIDAN AV SKRIVBLADEN

Markera besvarade uppgifter med kryss

1	2	3	4	5	6	7	8	9	Antal inl. blad
X	X	X	X	X					7
Lär.ant. 20p	19p	14p	20p	20p					

Gk

POÄNG	BETYG	Lärarens sign.
93p	A	RC

SU, STATISTIK

Skrivsal: Brunnsvikssalen Anonymkod: [redacted] Blad nr: 1

1) 20p

Oberoende observationer

(1)

$n = 4$

$\mu = ?$

$\sigma^2 = ?$

$\bar{x} = 80$ $\qquad \frac{78 + 80 + 82 + 80}{4} = \frac{320}{4} = 80$

$s = \sqrt{8/3}$

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2$$

$$= \frac{1}{3} \left((78-80)^2 + (80-80)^2 + (82-80)^2 + (80-80)^2 \right)$$

$$= \frac{1}{3} (4 + 0 + 4 + 0)$$

$$= \frac{1}{3} \cdot 8 = \frac{8}{3}$$

$$s = \sqrt{s^2} = \sqrt{\frac{8}{3}}$$

NÄSTA SIDA! → →

1 a)

Konfidensintervall = KI

Frihetsgrader $n-1 = 4-1 = 3$

95 % KI : $t_{\alpha/2,\, n-1} \Leftrightarrow \boxed{t_{0{,}025,\,3} = 3.18 \text{ (enligt tabell)}}$

$\alpha/2 = 0{,}025$

Pga väntevärde och varians okända observationerna kan uppfattas som oberoende gör vi en normalfördelning

Beräkning av KI

$$\bar{x} \pm t \cdot s/\sqrt{n}$$

$$80 \pm 3{,}18 \cdot \frac{\sqrt{8/3}}{\sqrt{4}} =$$

$$= 80 \pm 2{,}6 \Rightarrow [77{,}4 ; 82{,}6]$$

Slutsats: I 95 fall av 100 kommer den genomsnittliga vikten för populationen ligga inom intervallet

[77,4 ; 82,6]

SU, STATISTIK

Skrivsal: Brumsvikshallen Anonymkod: [redacted] Blad nr: 2

1 b) Hypoteser: $H_0: \mu \leq 80$

$H_A: \mu > 80$

Testvariabel

Ty stickprovet litet, varians och väntevärde okänd samt observationerna obesoende gör vi en normalfördelning med t som testvariabel.

$$t = \frac{\bar{X} - \mu_0}{s/\sqrt{n}}$$

signifikansnivå = 5%. $t_{\alpha, n-1} = t_{0,05, 3}$

Beslutsregel

(enligt tabell)

Förkasta H_0 om $t > 2,35$

Beräkning

$$t = \frac{\bar{X} - \mu_0}{s/\sqrt{n}} = \frac{80 - 80}{\sqrt{\frac{8}{3}}/\sqrt{4}} = \frac{0}{\sqrt{\frac{8}{3}}/2} = 0$$

$0 < 2,35$

Slutsats:

Eftersom $0 < 2,35$ kan vi EJ förkasta H_0 på signifikansnivå 5%. μ kan anta ett mindre värde än 80

SU, STATISTIK

Skrivsal: Brunnsvikssalen Anonymkod [redacted] Blad nr: 3

2) 19p

2)

a) $n = 400$

$\hat{p} = \frac{400}{500} = 0.8$

$P_0 = 0,75$

Ty stickprovet är stort säger centrala gränsvärdessatsen att vi kan anta en normalfördelning med Z som testvariabel

Signifikansnivå: 5%

$Z_\alpha : Z_{0.05} = 1,6449$ (enligt tabell)

Hypoteser

$H_0 : P \leq 0.75$

$H_A : P > 0,75$

Testvariabel

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{\hat{p} \cdot (1-\hat{p})}{n}}}$$

Beslutsregel

Förkasta H_0 om $Z_{...} > Z_{0.05}$ $\boxed{1,6449 \text{ (enligt tabell)}}$

R

Beräkning:

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{\hat{p} \cdot (1-\hat{p})}{n}}} = \frac{0,8 - 0,75}{\sqrt{\frac{0,8 \cdot (1-0,8)}{400}}} = \frac{0,05}{0,02} = 2,5$$

2,8

F

= 500

F

$2,5 > 1,6449$

Nästa sida → →

Beslut

Eftersom 2,5 > 1,6449 Förkastar vi H_0!
Andelen familjer som har köpt product X i
New York är större än 0,75 på signifikansnivå
5%.

b) p-värde = $1 - \Phi(2,5)$

$= 1 - 0,99506 = 0,00494$

Sannolikheten att H_0 inträffar för populationen
är 0,00494 vilket är väldigt mycket mindre än
den valda signifikansnivån.

SU, STATISTIK

Skrivsal: Brunnsviksalen   Anonymkod: [redacted] Blad nr: 4

3) ~~...~~ 14p

X   Y

3) $n = 10$ Grupp A. 5 st   Grupp B 5 st

$$\bar{X} = \frac{10,9 + 8,1 + 8,5 + 8,0 + 9,6}{5} = 9,02$$

12

$$\bar{Y} = \frac{9,7 + 5,7 + 8,2 + 6,9 + 8,4}{5} = 7,78$$

$\sigma_A^2 - \sigma_B^2 = 0$ (Dom är lika stora)

Blodsockernivåerna lika antal som oberoende observationer

Hypotes

$H_0: \mu_A - \mu_B = 0$ $H_A: \mu_A - \mu_B \neq 0$

F (−2p)

Signifikansnivå: 10% = $t_{\alpha, n-1} = t_{0,1, 9} = 1,38$ enligt tabell

$= 1,86$

Då vi har ett litet stickprov, okända väntevärden samt att de okända variansena är lika stora gör vi en normalfördelning med t som testvariabel

Testvariabel

$$t = \frac{\bar{X} - \bar{Y} - D_0}{\sqrt{S_p^2\left(\frac{1}{n} + \frac{1}{m}\right)}} \quad \text{där } S_p^2 = \frac{(n-1)\,S_1^2 + (m-1)\,S_2^2}{n+m-2}$$

S^2 för grupp A $= \frac{1}{n-1} \sum_{i=1}^{n=5} (x_i - \bar{x})^2$

$= \frac{1}{4}\Big((10,9 - 9,02)^2 + (8,1 - 9,02)^2 + (8,5 - 9,02)^2 + (8,0 - 9,02)^2$

$+ (9,6 - 9,02)^2 \Big)$

$= \frac{1}{4} (3,534 + 0,846 + 0,27 + 1,04 + 0,336)$

$= \frac{1}{4} \cdot 6,026 = \frac{6,026}{4} = 1,5065$ $\boxed{S_A^2 = 1,5065}$

S^2 för grupp B $= \frac{1}{n-1} \sum_{i=1}^{n=5} (Y_i - \bar{Y})^2$

$= \frac{1}{4}\Big((9,7 - 7,78)^2 + (5,7 - 7,78)^2 + (8,2 - 7,78)^2 + (6,9 - 7,78)^2$

$+ (8,4 - 7,78)^2$

$= \frac{1}{4} (3,686 + 4,326 + 0,176 + 0,774 + 0,384)$

$= \frac{1}{4} \cdot 9,346 = \frac{9,346}{4} = 2,336$ $\boxed{S_B^2 = 2,336}$

SU, STATISTIK

Skrivsal: Brunnsvikssalen Anonymkod: [redacted] Blad nr: 5

(3) forts. lösning

Beslut: Förkasta H_0 om $t > 1,78$

Beräkning:

$$t = \frac{\bar{X} - \bar{Y} - D_0}{\sqrt{S_p^2\left(\frac{1}{n} + \frac{1}{m}\right)}} \quad \text{där} \quad S_p^2 = \frac{(n-1)S_1^2 + (m-1)S_2^2}{n+m-2}$$

$$S_p^2 = \frac{(10-1)\,1,5065 + (5-1)\,2,336}{10+5-2} = \frac{13,5585 + 9,344}{13} =$$

$$S_p^2 = 22,9 \quad = 1,93$$

F (-2p)

$$t = \frac{9,02 - 7,78 - 0}{\sqrt{22,9\left(\frac{1}{10} + \frac{1}{5}\right)}} = \frac{1,24}{\sqrt{6,87}} = \frac{1,24}{2,62} = \boxed{0,473}$$

F (-2p)

$0,473 \not> 1,38$

F F

Beslut $0,473 < 1,38$ så vi kan ej förkasta H_0

Vi kan ej se om det finns någon skillnad i effekt på blodsockernivåerna mellan de båda behandlingarna på signifikansnivån 10%.

3) GER ~~10p~~ 14p

SU, STATISTIK

Skrivsal: Brunnsvikssalen Anonymkod [redacted] Blad nr: 6

4) 20p

(4)

(4) **Hypoteser** $n = 100$

H_0: Påståendet stämmer EJ

H_A = Påståendet stämmer

Signifikansnivå 10% =

Då $n = 100$ räknar jag varuhusägarens förväntningar i procent i antal

OBSERVERAT	Förväntat
0 artiklar : 80	60
1 artikel : 10	30
Minst 2 : 10	10

$$\chi^2 = \sum \frac{(O-E)^2}{E} = \frac{(80-60)^2}{60} + \frac{(10-30)^2}{30} + \frac{(10-10)^2}{10}$$

$$= 6{,}66 + 13{,}33 = 19{,}99$$

R

$$\chi^2_{\alpha\,(r-1)(c-1)} = \chi^2_{0,1,4}$$

frihetsgrader: $(3-1)(3-1) = 9-3-3+1 = 4$

Jag hittar ej i χ^2 tabellen talet för χ^2 med signifikansnivå 10% med frihetsgrader 4

Men! om man tittar mellan 0,95 och 0,05 i tabellen (på 4 frihetsgrader) ser man att talet inte kan vara större än 19.99

Beslut

$\chi^2_{0,1,4} < 19,99$ och vi förkastar EJ H_0

Ägarens påstående stämmer ej på signifikans-nivå 10%.

SU, STATISTIK

Skrivsal: Brunnsvikssalen Anonymkod: [redacted] Blad nr: 7

5) 20p

(5)

Produkter/efterfrågan	E1	E2	E3	E4	VINST
A	10 mkr	10 mkr	10 mkr	10 mkr	10 mkr
B	18 mkr	18 mkr	2 mkr	10 mkr	12 mkr
C	2 mkr	30 mkr	0	0	8 Mkr
D	10 mkr	25 mkr	2 mkr	2 mkr	9,75 Mkr

B bör dom välja!!

a) Laplacekritenet förklarar att sannolikheten är lika lika stor för alla efterfrågorna att inträffa, och med att vi har 4 olika typer av efterfrågor för varje produkt ger Laplacekriteriet att varje efterfrågan har $\frac{1}{4}$ att inträffa

Beräkning

A : $\frac{1}{4}\cdot 10 + \frac{1}{4}\cdot 10 + \frac{1}{4}\cdot 10 + \frac{1}{4}\cdot 10 = 10$ Mkr

B : $\frac{1}{4}\cdot 18 + \frac{1}{4}\cdot 18 + \frac{1}{4}\cdot 2 + \frac{1}{4}\cdot 10 = 12$ Mkr

C : $\frac{1}{4}\cdot 2 + \frac{1}{4}\cdot 30 + \frac{1}{4}\cdot 0 + \frac{1}{4}\cdot 0 = 8$ Mkr

D : $\frac{1}{4}\cdot 10 + \frac{1}{4}\cdot 25 + \frac{1}{4}\cdot 2 + \frac{1}{4}\cdot 2 = 9{,}75$ Mkr

Svar: Företagsledningen Bör välja produkt B b)

Produkt/ Efterfrågan	E1 0,2	E2 0,4	E3 0,3	E4 0,1	YINST
A	10 Mkr	10 Mkr	10 Mkr	10 Mkr	10 Mkr
B	18 Mkr	18 Mkr	2 Mkr	10 Mkr	12,4 Mkr
C	2 Mkr	30 Mkr	0	0	12,4 Mkr
D	10 Mkr	25 Mkr	2 Mkr	2 Mkr	12,8 Mkr

Produkt D bör företaget välja!

Beräkning

A : $0,2 \cdot 10 + 0,4 \cdot 10 + 0,3 \cdot 10 + 0,1 \cdot 10 = 10$ Mkr

B : $0,2 \cdot 18 + 0,4 \cdot 18 + 0,3 \cdot 2 + 0,1 \cdot 10 = 12,4$ Mkr

C : $0,2 \cdot 2 + 0,4 \cdot 30 + 0,3 \cdot 0 + 0,1 \cdot 0 = 12,4$ Mkr

D : $0,2 \cdot 10 + 0,4 \cdot 25 + 0,3 \cdot 2 + 0,1 \cdot 2 = 12,8$ Mkr

SVAR Med den nya sannolikheten ser vi att företagsledningen borde välja produkt D!

R

LÖSNINGAR / STATISTIKENS GRUNDER 2 / 2013-01-15

1a) $\bar{X} = 80$ $\quad S^2 = 2{,}67$ $\quad S = 1{,}63$ $\quad n = 4$

$\alpha = 0{,}05$

$$\bar{X} \pm t_{\alpha/2\,(n-1)} \cdot \frac{S}{\sqrt{n}}$$

$\alpha/2 = 0{,}025$

$$80 \pm 3{,}18 \cdot \frac{1{,}63}{\sqrt{4}} \qquad 80 \pm 2{,}59 \qquad [77{,}41 \,;\, 82{,}59]$$

1b) $H_0: \mu \leq 80 \qquad H_A: \mu > 80$

FÖRKASTA H_0 OM $t > t_{\alpha(n-1)}$

EFTERSOM

$$t = \frac{\bar{X} - 80}{S/\sqrt{n}} = \frac{80-80}{1{,}63/\sqrt{4}} = 0 \not> t_{0{,}05}(3) = 2{,}35$$

H_0 KAN EJ FÖRKASTAS

2a) $H_0: p \leq 0{,}75 \qquad H_1: p > 0{,}75 \qquad n = 500$

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}} = \frac{0{,}8 - 0{,}75}{\sqrt{\frac{0{,}16}{500}}} = 2{,}80$$

$\hat{p} = \frac{400}{500} = 0{,}8$

$\hat{p}(1-\hat{p}) = (0{,}8)(0{,}2) = 0{,}16$

$Z_\alpha = Z_{0{,}05} = 1{,}64$ FÖRKASTA H_0 OM $Z > Z_\alpha$

$2{,}80 > 1{,}64$

H_0 FÖRKASTAS

2b) $p = P(Z > 2{,}80) = 1 - \Phi(2{,}80) = 1 - 0{,}99744 =$

$= 0{,}00256$

3) $H_0: \mu_A = \mu_B \quad H_1: \mu_A \neq \mu_B$

$$t = \frac{\bar{X} - \bar{Y}}{\sqrt{S_p^2\left(\frac{1}{n}+\frac{1}{m}\right)}} = \frac{9{,}02 - 7{,}78}{\sqrt{1{,}93\left(\frac{1}{5}+\frac{1}{5}\right)}} = \underline{1{,}41}$$

$\bar{X} = 9{,}02 \quad S_X^2 = 1{,}51 \quad \bar{Y} = 7{,}78 \quad S_Y^2 = 2{,}34$

$$S_p^2 = \frac{4 \cdot 1{,}51 + 4 \cdot 2{,}34}{5+5-2} = \frac{15{,}4}{8} = \underline{1{,}93}$$

$$t_{\alpha/2\,(n+m-2)} = t_{0{,}05\,(8)} = \underline{1{,}86}$$

FÖRKASTA H_0 OM $t < -t_{0{,}05(8)}$ ELLER $t > t_{0{,}05(8)}$

$1{,}41 \not> 1{,}86$

VI KAN EJ FÖRKASTA H_0

4)

	O_i	E_i	$(O_i - E_i)^2/E_i$
0	80	60	6,6666
1	10	30	13,3333
MINST 2	10	10	0
	100	100	20

$\chi^2 = 20 > \chi^2_{0{,}05(2)} = 5{,}99$

FÖRKASTAS ÄGARENS PÅSTÅENDE

5)

	E_1	E_2	E_3	E_4	LAPLACE	GIVNA S.
A	10	10	10	10	10	10
B	18	18	2	10	12	12,4
C	2	30	0	0	8	12,4
D	10	25	2	2	9,75	12,8
	0,2	0,4	0,3	0,1		

5 a) ALTERNATIV B $(0{,}25 \times 18 + 0{,}25 \times 18 + 0{,}25 \times 2 + 0{,}25 \times 10 = 12)$

5b) ALT. D $(0{,}2 \times 10 + 0{,}4 \times 25 + 0{,}3 \times 2 + 0{,}1 \times 2 = 12{,}8)$