Stockholms universitet
Statistiska institutionen
Raul Cano

SKRIVNINGSDATUM: 2015-06-05

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

Skrivtid: 5 timmar
Hjälpmedel: Miniräknare. Vidhäftade formel- och tabellblad (obs! vidhäftas endast de tabellsidor som behövs för den här tentamen).
Tentamensgenomgång och återlämning: Tisdagen den 16 juni, 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)

Anta att vikten på Chiquita-bananer är normalfördelad med okänt väntevärde μ och okänd standardavvikelse σ. För 16 slumpmässigt utvalda Chiquita-bananer har man observerat ett medelvärde lika med 228 gram och en standardavvikelse lika med 5.

a) Beräkna ett 90 % konfidensintervall för μ. (5 p.)

b). Testa att $\mu < 230$ gram på risknivån (signifikansnivån) 5 %. Vilken blir din slutsats? (10 p.)

c). Beräkna (approximativt) p-värdet i samband med ovanstående hypotesprövning. Använd de tabeller som bifogas i tentan. (5 p.)

Uppgift 2: (20 poäng)

För en mycket stor population av personer anställda inom en viss bransch, gäller att en okänd andel p av dem har en inkomst mindre än 190 000 kronor. Av 100 slumpmässigt utvalda personer från populationen uppgav 47 att de har en inkomst mindre än 190 000 kronor.

a) Beräkna ett 95 % konfidensintervall för p. (5 p.)

b). Testa att den okända andelen p är större än 0,40 på risknivån 1 %. Vilken blir din slutsats? (10 p.)

c). Beräkna (approximativt) p-värdet i samband med ovanstående hypotesprövning. Använd de tabeller som bifogas i tentan. (5 p.)

Uppgift 3: (20 poäng)

Före ett presidentval uppgav 96 av 120 slumpmässigt utvalda röstberättigade i valdistrikt A att de tänkte rösta på en viss kandidat. Av 140 slumpmässigt utvalda röstberättigade från valdistrikt B uppgav 84 att de tänkte rösta på den aktuella kandidaten.

a). Testa på risknivån 1 % om det finns en skillnad mellan de två valdistrikten med avseende på andelen som (om de blev tillfrågade) skulle uppge att de tänkte rösta på kandidaten ifråga. Obs! Genomför en hypotesprövning inte ett konfidensintervall! (10 p.)

b). Ange vilka hypoteser du testar. (5 p.)

c). Vilken blir din slutsats? (5 p.)

Uppgift 4: (20 poäng)

Ett varuhus genomför en marknadsundersökning för ta reda på om manliga och kvinnliga kunder skiljer sig åt i fråga om preferenserna för solprodukter Nivea, L'Oréal Paris, Garnier och Hawaiian Tropic. Varje person i ett slumpmässigt stickprov av 500 kunder tillfrågades om vilken av de fyra solprodukterna han/hon föredrog. Resultatet redovisas i tabellen nedan. (Talet i varje tabellcell står för antalet intervjuade personer som föredrar produkten i fråga.)

Solprodukter	Män	Kvinnor
Nivea	50	75
L'Oréal Paris	25	100
Garnier	75	50
Hawaiian Tropic	50	75

a). Ställ upp hypoteser (5 p.) och

b). Testa på signifikansnivån 1 % om det i populationen finns något beroende mellan kön och vilken solprodukt som föredras. Vilken blir din slutsats? (15 p.)

Uppgift 5: (20 poäng)

Låt oss anta följande beslutsmatris (kapitel 19 i kompendiet):

	Naturtillstånden	
Handlingsalternativ	S1	S2
A1	15	20
A2	25	15
A3	8	13
A4	15	12
A5	30	5

a) Bestäm med hjälp av Laplacekriteriet vilket handlingsalternativ man bör välja. (10 p.)

b) Anta att sannolikheterna för att de möjliga naturtillstånden inträffar är 0.8 respektive 0.2 för S1 respektive S2 och bestäm vilket handlingsalternativ man bör välja. (10 p.)

FORMLER

VT2013

Räkneregler för väntevärden och varianser (a, b och c är konstanter och X och Y är stokastiska variabler)

$E(c) = c$	$V(c) = 0$
$E(X + c) = E(X) + c$	$V(X + c) = V(X)$
$E(aX) = aE(X)$	$V(aX) = a^2V(X)$
$E(aX + bY + c) = aE(X) + bE(Y) + c$	$V(aX + bY + c) = a^2V(X) + b^2V(Y) + 2abCov(X, Y)$

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

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

Stickprovskovarians: $s_{xy} = \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = \frac{1}{n-1}(\sum_{i=1}^{n} x_i y_i - n\bar{x}\bar{y})$

Binomialfördelningen: $f(x) = \binom{n}{x} p^x(1-p)^{n-x} = \frac{n!}{x!(n-x)!}p^x(1-p)^{n-x}$

Poissonfördelningen: $f(x) = \frac{\lambda^x e^{-\lambda}}{x!}$

Diverse konfidensintervall och enkelsidiga testvariabler (*f.g.* = frihetsgrader):

$$\bar{x} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \qquad Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \geq z_\alpha$$

$$\bar{x} \pm t_{\alpha/2}^{(f.g.)}\frac{s}{\sqrt{n}} \qquad T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \geq t_\alpha^{(f.g.)}$$

$$\bar{x} \pm z_{\alpha/2}\frac{s}{\sqrt{n}} \qquad Z = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \geq z_\alpha$$

$$\bar{x}_1 - \bar{x}_2 \pm z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} \qquad Z = \frac{\bar{X}_1 - \bar{X}_2 - 0}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}} \geq z_\alpha$$

$$\bar{x}_1 - \bar{x}_2 \pm t_{\alpha/2}^{(f.g.)} \cdot s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \qquad T = \frac{\bar{X}_1 - \bar{X}_2 - 0}{S_p\sqrt{1/n_1 + 1/n_2}} \geq t_\alpha^{(f.g.)}$$

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

$$\bar{x}_1 - \bar{x}_2 \pm z_{\alpha/2}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \qquad Z = \frac{\bar{X}_1 - \bar{X}_2 - 0}{\sqrt{S_1^2/n_1 + S_2^2/n_2}} \geq z_\alpha$$

Forts. konfidensintervall och enkelsidiga testvariabler (*f.g.* = frihetsgrader):

$$\bar{d} \pm t_{\alpha/2}^{(f.g.)} \frac{s_d}{\sqrt{n}} \qquad T = \frac{\bar{D} - 0}{S_D/\sqrt{n}} \geq t_{\alpha}^{(f.g.)}$$

$$\frac{y}{n} \pm z_{\alpha/2}\sqrt{(y/n)(1 - y/n)/n} \qquad Z = \frac{Y/n - \pi_0}{\sqrt{\pi_0(1 - \pi_0)/n}} \geq z_\alpha$$

$$\hat{p}_1 - \hat{p}_2 \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1 - \hat{p}_1)}{n_1} + \frac{\hat{p}_2(1 - \hat{p}_2)}{n_2}} \qquad Z = \frac{Y_1/n_1 - Y_2/n_2 - 0}{\sqrt{\left(\frac{Y_1+Y_2}{n_1+n_2}\right)\left(1 - \frac{Y_1+Y_2}{n_1+n_2}\right)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \geq z_\alpha$$

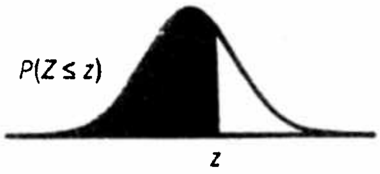
$$\chi^2 = \sum \frac{(n_i - n\pi_i)^2}{n\pi_i} \geq \chi_\alpha^2(f.g.)$$

$$\chi^2 = \sum\sum \frac{(n_{ij} - n_{i.}n_{.j}/n)^2}{n_{i.}n_{.j}/n} \geq \chi_\alpha^2(f.g.)$$

TABELL 1. Normalfördelningen, standardiserad

$\Phi(z) = P(Z \le z)$ där $Z \in N(0, 1)$.

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

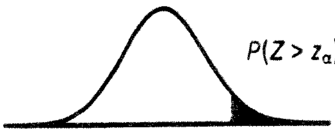


z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,50000	0,50399	0,50798	0,51197	0,51595	0,51994	0,52392	0,52790	0,53188	0,53586
0,1	0,53983	0,54380	0,54776	0,55172	0,55567	0,55962	0,56356	0,56749	0,57142	0,57535
0,2	0,57926	0,58317	0,58706	0,59095	0,59483	0,59871	0,60257	0,60642	0,61026	0,61409
0,3	0,61791	0,62172	0,62552	0,62930	0,63307	0,63683	0,64058	0,64431	0,64803	0,65173
0,4	0,65542	0,65910	0,66276	0,66640	0,67003	0,67364	0,67724	0,68082	0,68439	0,68793
0,5	0,69146	0,69497	0,69847	0,70194	0,70540	0,70884	0,71226	0,71566	0,71904	0,72240
0,6	0,72575	0,72907	0,73237	0,73565	0,73891	0,74215	0,74537	0,74857	0,75175	0,75490
0,7	0,75804	0,76115	0,76424	0,76730	0,77035	0,77337	0,77637	0,77935	0,78230	0,78524
0,8	0,78814	0,79103	0,79389	0,79673	0,79955	0,80234	0,80511	0,80785	0,81057	0,81327
0,9	0,81594	0,81859	0,82121	0,82381	0,82639	0,82894	0,83147	0,83398	0,83646	0,83891
1,0	0,84134	0,84375	0,84614	0,84849	0,85083	0,85314	0,85543	0,85769	0,85993	0,86214
1,1	0,86433	0,86650	0,86864	0,87076	0,87286	0,87493	0,87698	0,87900	0,88100	0,88298
1,2	0,88493	0,88686	0,88877	0,89065	0,89251	0,89435	0,89617	0,89796	0,89973	0,90147
1,3	0,90320	0,90490	0,90658	0,90824	0,90988	0,91149	0,91309	0,91466	0,91621	0,91774
1,4	0,91924	0,92073	0,92220	0,92364	0,92507	0,92647	0,92785	0,92922	0,93056	0,93189
1,5	0,93319	0,93448	0,93574	0,93699	0,93822	0,93943	0,94062	0,94179	0,94295	0,94408
1,6	0,94520	0,94630	0,94738	0,94845	0,94950	0,95053	0,95154	0,95254	0,95352	0,95449
1,7	0,95543	0,95637	0,95728	0,95818	0,95907	0,95994	0,96080	0,96164	0,96246	0,96327
1,8	0,96407	0,96485	0,96562	0,96638	0,96712	0,96784	0,96856	0,96926	0,96995	0,97062
1,9	0,97128	0,97193	0,97257	0,97320	0,97381	0,97441	0,97500	0,97558	0,97615	0,97670
2,0	0,97725	0,97778	0,97831	0,97882	0,97932	0,97982	0,98030	0,98077	0,98124	0,98169
2,1	0,98214	0,98257	0,98300	0,98341	0,98382	0,98422	0,98461	0,98500	0,98537	0,98574
2,2	0,98610	0,98645	0,98679	0,98713	0,98745	0,98778	0,98809	0,98840	0,98870	0,98899
2,3	0,98928	0,98956	0,98983	0,99010	0,99036	0,99061	0,99086	0,99111	0,99134	0,99158
2,4	0,99180	0,99202	0,99224	0,99245	0,99266	0,99286	0,99305	0,99324	0,99343	0,99361
2,5	0,99379	0,99396	0,99413	0,99430	0,99446	0,99461	0,99477	0,99492	0,99506	0,99520
2,6	0,99534	0,99547	0,99560	0,99573	0,99585	0,99598	0,99609	0,99621	0,99632	0,99643
2,7	0,99653	0,99664	0,99674	0,99683	0,99693	0,99702	0,99711	0,99720	0,99728	0,99736
2,8	0,99744	0,99752	0,99760	0,99767	0,99774	0,99781	0,99788	0,99795	0,99801	0,99807
2,9	0,99813	0,99819	0,99825	0,99831	0,99836	0,99841	0,99846	0,99851	0,99856	0,99861
3,0	0,99865	0,99869	0,99874	0,99878	0,99882	0,99886	0,99889	0,99893	0,99896	0,99900
3,1	0,99903	0,99906	0,99910	0,99913	0,99916	0,99918	0,99921	0,99924	0,99926	0,99929
3,2	0,99931	0,99934	0,99936	0,99938	0,99940	0,99942	0,99944	0,99946	0,99948	0,99950
3,3	0,99952	0,99953	0,99955	0,99957	0,99958	0,99960	0,99961	0,99962	0,99964	0,99965
3,4	0,99966	0,99968	0,99969	0,99970	0,99971	0,99972	0,99973	0,99974	0,99975	0,99976
3,5	0,99977	0,99978	0,99978	0,99979	0,99980	0,99981	0,99981	0,99982	0,99983	0,99983
3,6	0,99984	0,99985	0,99985	0,99986	0,99986	0,99987	0,99987	0,99988	0,99988	0,99989
3,7	0,99989	0,99990	0,99990	0,99990	0,99991	0,99991	0,99992	0,99992	0,99992	0,99992
3,8	0,99993	0,99993	0,99993	0,99994	0,99994	0,99994	0,99994	0,99995	0,99995	0,99995
3,9	0,99995	0,99995	0,99996	0,99996	0,99996	0,99996	0,99996	0,99996	0,99997	0,99997
4,0	0,99997	0,99997	0,99997	0,99997	0,99997	0,99997	0,99998	0,99998	0,99998	0,99998

TABELL 2. Normalfördelningens kvantiler, standardiserad

$Z \in N(0, 1)$. Vilket värde har z_α om $P(Z > z_\alpha) = \alpha$ där α är en given sannolikhet.

Utnyttja även $\Phi(-z) = 1 - \Phi(z)$ för $P(Z \leq -z_\alpha)$.

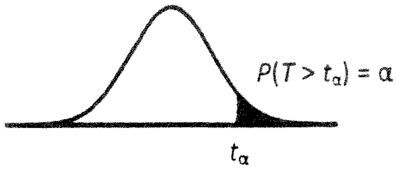


α	z_α
0,1	1,2816
0,05	1,6449
0,025	1,9600
0,010	2,3263
0,005	2,5758
0,0025	2,8070
0,0010	3,0902
0,0005	3,2905
0,00025	3,4808
0,00010	3,7190
0,00005	3,8906
0,000025	4,0556
0,000010	4,2649
0,000005	4,4172

TABELL 3. t-fördelningens kvantiler

$T \in t(v)$ där v = antal frihetsgrader.

Vilket värde har t_α om $P(T > t_\alpha) = \alpha$ där α är en given sannolikhet. Utnyttja även $P(T \leq -t_\alpha) = P(T > t_\alpha)$.



v	α = 0,1	0,05	0,025	0,010	0,005	0,0025	0,0010	0,0005
1	3,078	6,314	12,706	31,821	63,657	127,321	318,309	636,619
2	1,886	2,920	4,303	6,965	9,925	14,089	22,327	31,599
3	1,638	2,353	3,182	4,541	5,841	7,453	10,215	12,924
4	1,533	2,132	2,776	3,747	4,604	5,598	7,173	8,610
5	1,476	2,015	2,571	3,365	4,032	4,773	5,893	6,869
6	1,440	1,943	2,447	3,143	3,707	4,317	5,208	5,959
7	1,415	1,895	2,365	2,998	3,499	4,029	4,785	5,408
8	1,397	1,860	2,306	2,896	3,355	3,833	4,501	5,041
9	1,383	1,833	2,262	2,821	3,250	3,690	4,297	4,781
10	1,372	1,812	2,228	2,764	3,169	3,581	4,144	4,587
11	1,363	1,796	2,201	2,718	3,106	3,497	4,025	4,437
12	1,356	1,782	2,179	2,681	3,055	3,428	3,930	4,318
13	1,350	1,771	2,160	2,650	3,012	3,372	3,852	4,221
14	1,345	1,761	2,145	2,624	2,977	3,326	3,787	4,140
15	1,341	1,753	2,131	2,602	2,947	3,286	3,733	4,073
16	1,337	1,746	2,120	2,583	2,921	3,252	3,686	4,015
17	1,333	1,740	2,110	2,567	2,898	3,222	3,646	3,965
18	1,330	1,734	2,101	2,552	2,878	3,197	3,610	3,922
19	1,328	1,729	2,093	2,539	2,861	3,174	3,579	3,883
20	1,325	1,725	2,086	2,528	2,845	3,153	3,552	3,850
21	1,323	1,721	2,080	2,518	2,831	3,135	3,527	3,819
22	1,321	1,717	2,074	2,508	2,819	3,119	3,505	3,792
23	1,319	1,714	2,069	2,500	2,807	3,104	3,485	3,768
24	1,318	1,711	2,064	2,492	2,797	3,091	3,467	3,745
25	1,316	1,708	2,060	2,485	2,787	3,078	3,450	3,725
26	1,315	1,706	2,056	2,479	2,779	3,067	3,435	3,707
27	1,314	1,703	2,052	2,473	2,771	3,057	3,421	3,690
28	1,313	1,701	2,048	2,467	2,763	3,047	3,408	3,674
29	1,311	1,699	2,045	2,462	2,756	3,038	3,396	3,659
30	1,310	1,697	2,042	2,457	2,750	3,030	3,385	3,646
35	1,306	1,690	2,030	2,438	2,724	2,996	3,340	3,591
40	1,303	1,684	2,021	2,423	2,704	2,971	3,307	3,551
45	1,301	1,679	2,014	2,412	2,690	2,952	3,281	3,520
50	1,299	1,676	2,009	2,403	2,678	2,937	3,261	3,496
55	1,297	1,673	2,004	2,396	2,668	2,925	3,245	3,476
60	1,296	1,671	2,000	2,390	2,660	2,915	3,232	3,460
65	1,295	1,669	1,997	2,385	2,654	2,906	3,220	3,447
70	1,294	1,667	1,994	2,381	2,648	2,899	3,211	3,435
75	1,293	1,665	1,992	2,377	2,643	2,892	3,202	3,425

Forts. nästa sida

TABELL 3 forts. *t*-fördelningens kvantiler

v	α = 0,1	0,05	0,025	0,010	0,005	0,0025	0,0010	0,0005
80	1,292	1,664	1,990	2,374	2,639	2,887	3,195	3,416
85	1,292	1,663	1,988	2,371	2,635	2,882	3,189	3,409
90	1,291	1,662	1,987	2,368	2,632	2,878	3,183	3,402
95	1,291	1,661	1,985	2,366	2,629	2,874	3,178	3,396
100	1,290	1,660	1,984	2,364	2,626	2,871	3,174	3,390
125	1,288	1,657	1,979	2,357	2,616	2,858	3,157	3,370
150	1,287	1,655	1,976	2,351	2,609	2,849	3,145	3,357
175	1,286	1,654	1,974	2,348	2,604	2,843	3,137	3,347
200	1,286	1,653	1,972	2,345	2,601	2,839	3,131	3,340
300	1,284	1,650	1,968	2,339	2,592	2,828	3,118	3,323
400	1,284	1,649	1,966	2,336	2,588	2,823	3,111	3,315
500	1,283	1,648	1,965	2,334	2,586	2,820	3,107	3,310
1000	1,282	1,646	1,962	2,330	2,581	2,813	3,098	3,300
2000	1,282	1,646	1,961	2,328	2,578	2,810	3,094	3,295
3000	1,282	1,645	1,961	2,328	2,577	2,809	3,093	3,294
4000	1,282	1,645	1,961	2,327	2,577	2,809	3,092	3,293
5000	1,282	1,645	1,960	2,327	2,577	2,808	3,092	3,292

TABELL 4. χ^2-fördelningens kvantiler

$Q \in \chi^2(v)$ där v = antal frihetsgrader.

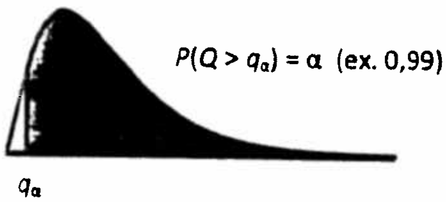
Vilket värde har q_α om $P(Q > q_\alpha) = \alpha$ där α är en sannolikhet.

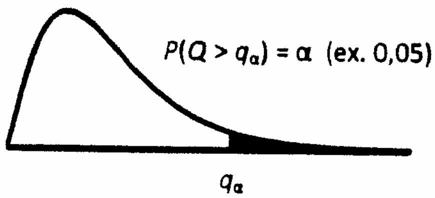
v	α = 0,999	0,995	0,99	0,975	0,95	0,05	0,025	0,01	0,005	0,001
1	0,000	0,000	0,000	0,001	0,004	3,841	5,024	6,635	7,879	10,828
2	0,002	0,010	0,020	0,051	0,103	5,991	7,378	9,210	10,597	13,816
3	0,024	0,072	0,115	0,216	0,352	7,815	9,348	11,345	12,838	16,266
4	0,091	0,207	0,297	0,484	0,711	9,488	11,143	13,277	14,860	18,467
5	0,210	0,412	0,554	0,831	1,145	11,070	12,833	15,086	16,750	20,515
6	0,381	0,676	0,872	1,237	1,635	12,592	14,449	16,812	18,548	22,458
7	0,598	0,989	1,239	1,690	2,167	14,067	16,013	18,475	20,278	24,322
8	0,857	1,344	1,646	2,180	2,733	15,507	17,535	20,090	21,955	26,124
9	1,152	1,735	2,088	2,700	3,325	16,919	19,023	21,666	23,589	27,877
10	1,479	2,156	2,558	3,247	3,940	18,307	20,483	23,209	25,188	29,588
11	1,834	2,603	3,053	3,816	4,575	19,675	21,920	24,725	26,757	31,264
12	2,214	3,074	3,571	4,404	5,226	21,026	23,337	26,217	28,300	32,909
13	2,617	3,565	4,107	5,009	5,892	22,362	24,736	27,688	29,819	34,528
14	3,041	4,075	4,660	5,629	6,571	23,685	26,119	29,141	31,319	36,123
15	3,483	4,601	5,229	6,262	7,261	24,996	27,488	30,578	32,801	37,697
16	3,942	5,142	5,812	6,908	7,962	26,296	28,845	32,000	34,267	39,252
17	4,416	5,697	6,408	7,564	8,672	27,587	30,191	33,409	35,718	40,790
18	4,905	6,265	7,015	8,231	9,390	28,869	31,526	34,805	37,156	42,312
19	5,407	6,844	7,633	8,907	10,117	30,144	32,852	36,191	38,582	43,820
20	5,921	7,434	8,260	9,591	10,851	31,410	34,170	37,566	39,997	45,315
21	6,447	8,034	8,897	10,283	11,591	32,671	35,479	38,932	41,401	46,797
22	6,983	8,643	9,542	10,982	12,338	33,924	36,781	40,289	42,796	48,268
23	7,529	9,260	10,196	11,689	13,091	35,172	38,076	41,638	44,181	49,728
24	8,085	9,886	10,856	12,401	13,848	36,415	39,364	42,980	45,559	51,179
25	8,649	10,520	11,524	13,120	14,611	37,652	40,646	44,314	46,928	52,620
26	9,222	11,160	12,198	13,844	15,379	38,885	41,923	45,642	48,290	54,052
27	9,803	11,808	12,879	14,573	16,151	40,113	43,195	46,963	49,645	55,476
28	10,391	12,461	13,565	15,308	16,928	41,337	44,461	48,278	50,993	56,892
29	10,986	13,121	14,256	16,047	17,708	42,557	45,722	49,588	52,336	58,301
30	11,588	13,787	14,953	16,791	18,493	43,773	46,979	50,892	53,672	59,703
32	12,811	15,134	16,362	18,291	20,072	46,194	49,480	53,486	56,328	62,487
34	14,057	16,501	17,789	19,806	21,664	48,602	51,966	56,061	58,964	65,247
36	15,324	17,887	19,233	21,336	23,269	50,998	54,437	58,619	61,581	67,985
38	16,611	19,289	20,691	22,878	24,884	53,384	56,896	61,162	64,181	70,703
40	17,916	20,707	22,164	24,433	26,509	55,758	59,342	63,691	66,766	73,402
42	19,239	22,138	23,650	25,999	28,144	58,124	61,777	66,206	69,336	76,084
44	20,576	23,584	25,148	27,575	29,787	60,481	64,201	68,710	71,893	78,750
46	21,929	25,041	26,657	29,160	31,439	62,830	66,617	71,201	74,437	81,400
48	23,295	26,511	28,177	30,755	33,098	65,171	69,023	73,683	76,969	84,037
50	24,674	27,991	29,707	32,357	34,764	67,505	71,420	76,154	79,490	86,661

Forts. nästa sida

TABELL 4 forts. χ^2-fördelningens kvantiler

v	α = 0,999	0,995	0,99	0,975	0,95	0,05	0,025	0,01	0,005	0,001
55	28,173	31,735	33,570	36,398	38,958	73,311	77,380	82,292	85,749	93,168
60	31,738	35,534	37,485	40,482	43,188	79,082	83,298	88,379	91,952	99,607
65	35,362	39,383	41,444	44,603	47,450	84,821	89,177	94,422	98,105	105,988
70	39,036	43,275	45,442	48,758	51,739	90,531	95,023	100,425	104,215	112,317
75	42,757	47,206	49,475	52,942	56,054	96,217	100,839	106,393	110,286	118,599
80	46,520	51,172	53,540	57,153	60,391	101,879	106,629	112,329	116,321	124,839
85	50,320	55,170	57,634	61,389	64,749	107,522	112,393	118,236	122,325	131,041
90	54,155	59,196	61,754	65,647	69,126	113,145	118,136	124,116	128,299	137,208
95	58,022	63,250	65,898	69,925	73,520	118,752	123,858	129,973	134,247	143,344
100	61,918	67,328	70,065	74,222	77,929	124,342	129,561	135,807	140,169	149,449
120	77,755	83,852	86,923	91,573	95,705	146,567	152,211	158,950	163,648	173,617
150	102,113	109,142	112,668	117,985	122,692	179,581	185,800	193,208	198,360	209,265
200	143,843	152,241	156,432	162,728	168,279	233,994	241,058	249,445	255,264	267,541
300	229,963	240,663	245,972	253,912	260,878	341,395	349,874	359,906	366,844	381,425
400	318,260	330,903	337,155	346,482	354,641	447,632	457,305	468,724	476,606	493,132
500	407,947	422,303	429,388	439,936	449,147	553,127	563,852	576,493	585,207	603,446





LÖSNINGAR / STATISTIKENS GRUNDER 2 / 2015-06-05

1) $X \sim N(\mu, \sigma)$, $n = 16$, $\bar{x} = 228$ gr, $s = 5$

a) $\bar{x} \pm t_{\alpha/2(n-1)} \cdot \frac{s}{\sqrt{n}}$ $\quad \alpha = 0{,}10$, $\alpha/2 = 0{,}05$ $\quad t_{0{,}05(15)} = 1{,}753$

$228 \pm 1{,}753 \cdot \frac{5}{4}$ $\quad 228 \pm 2{,}19125$ $\quad [225{,}80 ; 230{,}19]$

b) $H_0: \mu \geq 230$ $\quad H_1: \mu < 230$

FÖRKASTA H_0 OM $t < -t_{\alpha(n-1)}$ $\quad \alpha = 0{,}05$ $\quad -t_{0{,}05(15)} = -1{,}753$

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{228 - 230}{5/\sqrt{16}} = -\frac{2}{5/4} = -\frac{4}{5}(2) = -\frac{8}{5} = -1{,}6$$

EFTERSOM $-1{,}6 \not< -1{,}753$ H_0 EJ FÖRKASTAS.

c) P-VÄRDET $= P(t_{(n-1)} > |t_{OBS}| \mid H_0 \text{ SANN}) = P(t_{(15)} > 1{,}6)$

ENLIGT TABELL 3 LIGGER DET MELLAN 0,10 OCH 0,05
UNGEFÄR 0,07 (APPROXIMATIVT).

---

2) $\hat{p} \pm Z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ $\quad \hat{p} = \frac{47}{100} = 0{,}47$ $\quad \alpha = 0{,}05$ $\quad \frac{\alpha}{2} = 0{,}025$

a) $Z_{0{,}025} = 1{,}96$

$0{,}47 \pm 1{,}96 \cdot \sqrt{\frac{(0{,}47)(0{,}53)}{100}}$ $\quad 0{,}47 \pm 1{,}96 \sqrt{0{,}002491}$

$0{,}47 \pm 0{,}097823$ $\quad [0{,}37217 ; 0{,}56782]$

b) $H_0: p \leq 0{,}40$ $\quad H_1: p > 0{,}40$ $\quad \alpha = 0{,}01$ $\quad Z_\alpha = 2{,}3263$

FÖRKASTA H_0 OM $Z > Z_\alpha$

$$Z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}} = \frac{0{,}47 - 0{,}40}{\sqrt{\frac{(0{,}4)(0{,}6)}{100}}} = \frac{0{,}07}{0{,}048989795} = 1{,}428869$$

EFTERSOM $1{,}428869 \not> 2{,}3263$ H_0 EJ FÖRKASTAS.

c) $P(Z > 1{,}428869)$ = P-VÄRDET, ENLIGT TABELL 2
LIGGER MELLAN 0,10 OCH 0,05; $\approx 0{,}07$ (APPROXIMATIVT).

3) b) $H_0: P_A = P_B \quad H_1: P_A \neq P_B$

a) $\hat{P}_A = \frac{96}{120} = 0{,}8 \quad \hat{P}_B = \frac{84}{140} = 0{,}6 \quad \alpha = 0{,}01 \quad Z_{\alpha/2} = 2{,}5758$

$\hat{P}_{A+B} = \frac{96+84}{120+140} = \frac{180}{260} = 0{,}6923 \quad 1-\hat{P}_{A+B} = 0{,}3077$

$$\boxed{\frac{1}{n_1} + \frac{1}{n_2} = \frac{1}{120} + \frac{1}{140} = \frac{260}{16800} = 0{,}01547}$$

$$Z = \frac{0{,}8-0{,}6}{\sqrt{(0{,}6923)(0{,}3077)(0{,}01547)}} = \frac{0{,}2}{0{,}0574058394} = 3{,}483966$$

FÖRKASTA H_0 OM $Z < -Z_{\alpha/2}$ ELLER $Z > Z_{\alpha/2}$

EFTERSOM $Z = 3{,}483966 > Z_{\alpha/2} = 2{,}5758$ FÖRKASTAS H_0

c) DET FINNS EN SKILLNAD MELLAN DE TVÅ VALDISTRIKTEN MED AVSEENDE PÅ ANDELEN SOM (OM DE BLEV TILLFRÅGADE) SKULLE UPPGE ATT DE TÄNKTE RÖSTA PÅ KANDIDATEN I FRÅGA ($\alpha=0{,}01$).

4) a) H_0: SOLPRODUKTERS PREFERENSER ÄR OBEROENDE AV KÖN.

H_1: SOLPRODUKTERS PREFERENSER ÄR BEROENDE AV KÖN.

b)

50 (50)	75 (75)	125
25 (50)	100 (75)	125
75 (50)	50 (75)	125
50 (50)	75 (75)	125
200	300	500

FÖRKASTA H_0 OM $\chi^2 > \chi^2_{0{,}01(3)}$

$$\chi^2_{0{,}01(3)} = 11{,}345$$

$$\chi^2 = 0 + \frac{25^2}{50} + \frac{25^2}{50} + 0 + 0 + \frac{25^2}{75} + \frac{25^2}{75} + 0 = 41{,}66666$$

FÖRKASTA Ho.

5) a) LAPLACE

A1 $\quad 0{,}5(15) + 0{,}5(20) = 17{,}5$

A2 $\quad 0{,}5(25) + 0{,}5(15) = 20$ → VÄLJ A2

A3 $\quad 0{,}5(8) + 0{,}5(13) = 10{,}5$

A4 $\quad 0{,}5(15) + 0{,}5(12) = 13{,}5$

A5 $\quad 0{,}5(30) + 0{,}5(5) = 17{,}5$

b) A1 $\quad 0{,}8(15) + 0{,}2(20) = 16$

A2 $\quad 0{,}8(25) + 0{,}2(15) = 23$

A3 $\quad 0{,}8(8) + 0{,}2(13) = 9$

A4 $\quad 0{,}8(15) + 0{,}2(12) = 14{,}4$

A5 $\quad 0{,}8(30) + 0{,}2(5) = 25$ → VÄLJ A5

(13)

Statistiska institutionen

Stockholms universitet

Rättningsblad

Datum: 5/6 - 2015

Sal: Värtasalen

Tenta: Statistikens grunder 2 (kväll)

Kurs: Statistikens grunder

ANONYMKOD: SGK-0014

☒ Jag godkänner att min tenta får läggas ut anonymt på hemsidan som studentsvar.

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					5
Lär.ant. 20p	20p	20p	20p	20p					

POÄNG	BETYG	Lärarens sign.
100	A	RC

SU, STATISTIK

Skrivsal: VÄRTASALEN Anonymkod: SGK-0014 Blad nr: 1

1) 20 p

UPPGIFT 1

$X \sim N(\mu, \sigma)$

$n: 16$ $\bar{x}: 228$ $s: 5$

A

90% konfidensintervall för μ

Vi använder formeln nedan eftersom $n < 30$ och σ^2 är okänd

$$\bar{x} \pm t_{\frac{\alpha}{2}[n-1]} \frac{s}{\sqrt{n}}$$

Enligt tabell 3 ser vi att

$$t_{\frac{\alpha}{2}[n-1]} = t_{\frac{0,10}{2}[16-1]} = t_{0,05[15]} = 1,753$$

Vi har därmed alla element som behövs;

R

$$228 \pm 1,753 \cdot \frac{5}{\sqrt{16}}$$

$$228 \pm 1,753 \times 1,25$$

$$228 \pm 2,19125 \quad [225,80875 ; 230,19125]$$

R

Vi finner alltså att det 90%-iga konfidensintervallet har en lägsta punkt $\approx 225,8$

högsta punkt $\approx 230,2$

V. G. V. →

B

Testa $\mu < 230$ på risknivån 5%

$H_0: \mu \geq 230$ $\quad H_1: \mu < 230$

Beslutsregel H_0 förkastas om $t_{obs} < -t_{\alpha[n-1]}$

Enligt Tabell 3 $t_{\alpha[n-1]} = t_{0{,}05[15]} = 1{,}753$

t eftersom återigen $n < 30$ & okänd varians.

Testvariabel:

$$T = \frac{\bar{x} - \mu_0}{\frac{s}{\sqrt{n}}}$$

Vi har alla element & resultatet blir därför.

R

$$T = \frac{228 - 230}{\frac{5}{\sqrt{16}}} = \frac{-2}{1{,}25} = -1{,}6$$

R

Slutsats H_0 förkastas ej på 5% risknivå

eftersom $t_{obs} \cancel{<} -t_{\alpha[n-1]}$; $-1{,}6 > -1{,}753$

Vi kan inte utesluta att μ kan vara ≥ 230.

R

C

Beräkna approximativt p-värde

$P(t < -1{,}6) \rightarrow \quad P(t > 1{,}6)$

Enligt Tabell 3 ligger 1,6 mellan 0,1 och 0,05

R

Vilket indikerar att p-värdet ligger mellan

just 0,1 & 0,05

SU, STATISTIK

Skrivsal: VÄRTASALEN Anonymkod: SGK-0014 Blad nr: 2

2) 20p

A

UPPGIFT 2

$n: 100$ inkomst $< 190'$: 47 personer

Beräkna 95% konfidensintervall för p
Eftersom $n > 30$ använder vi
formeln/testvariabeln:

$$\frac{Y}{n} \pm Z_{\frac{\alpha}{2}} \sqrt{\left(\frac{Y}{n}\right)\left(1-\frac{Y}{n}\right)/n}$$

Vi vet att $y = 47$ $n = 100$ $\alpha = 0{,}05$
Enligt Tabell 2 ser vi att

$$Z_{\frac{\alpha}{2}} = Z_{\frac{0{,}05}{2}} = Z_{0{,}025} = 1{,}96$$ R

Nu har vi alla element som vi behövs:

$$\frac{47}{100} \pm 1{,}96 \sqrt{\left(\frac{47}{100}\right)\left(1-\frac{47}{100}\right)/100}$$

R

$$0{,}47 \pm 1{,}96 \times 0{,}0499$$

R R

$$0{,}47 \pm 0{,}0978 = [0{,}3722;\ 0{,}5678]$$

Vi finner alltså att det 95%-iga konfidensintervallet
R har lägsta punkt $\approx 0{,}37$
högsta punkt $\approx 0{,}57$

✓ $\hat{p}$ antas alltså ligga mellan detta intervall (95%).

V.G.V →

B

Testa $p > 0{,}40$ på risknivån 1%

$H_0 : p \leq 0{,}40$ $H_1 : p > 0{,}40$

Beslutsregel H_0 förkastas om $Z_{obs} > Z_\alpha$

$\alpha = 0{,}01$ och vi ser från Tabell 2 att

$Z_\alpha = Z_{0,01} = 2{,}3263$

Vi använder testvariabeln:

$$Z = \frac{\frac{Y}{n} - \pi_0}{\sqrt{\pi_0(1-\pi_0)/n}}$$

Vi har alla element som behövs och resultatet blir

$$Z = \frac{\frac{47}{100} - 0{,}40}{\sqrt{0{,}40(1-0{,}40)/100}} \approx \frac{0{,}07}{0{,}04899} = 1{,}42886$$

R

R

Slutsats: H_0 förkastas ej på 1% risknivå
eftersom $Z_{obs} \not> Z_{0,01}$, $1{,}42886 < 2{,}3263$

R

P kan alltså anta ett värde mindre eller lika med 0,40

C

Beräkna approximativt p-värde

R

$p(Z > 1{,}42886) \rightarrow 1 - p(Z \leq 1{,}42886)$

$\approx 1 - \Phi(1{,}43) = 1 - 0{,}92364 = 0{,}07636$

Approximativt p-värde

SU, STATISTIK

Skrivsal: VÄRTASALEN Anonymkod: SGK-0014 Blad nr: 3

3) 20p

(vill gärna sätta upp hypoteser först)

UPPGIFT 3

Valdistrikt A: n_1: 120 $Y_1 = 96$

Valdistrikt B: n_2: 140 $Y_2 = 84$

Risknivå 1% → α: 0,01

Vi ser från Tabell 2 att $Z_{\frac{0,01}{2}} = Z_{0,005} = 2,5758$

B

R

H_0: Ingen skillnad (mellan de två distrikten*) H_1: Skillnad (mellan de två distrikten*)

Beslutsregel: H_0 förkastas om $Z_{OBS} > Z_{\frac{\alpha}{2}}$

A

Testvariabel som vi använder:

$$Z = \frac{\frac{Y_1}{n_1} - \frac{Y_2}{n_2} - 0}{\sqrt{\left(\frac{Y_1+Y_2}{n_1+n_2}\right)\left(1-\frac{Y_1+Y_2}{n_1+n_2}\right)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$$

Vi har alla element som behövs och resultatet blir

$$Z = \frac{\frac{96}{120} - \frac{84}{140} - 0}{\sqrt{\left(\frac{96+84}{120+140}\right)\left(1-\frac{96+84}{120+140}\right)\left(\frac{1}{120}+\frac{1}{140}\right)}} \approx \frac{0,2}{0,05742} \approx 3,4831$$

R

C

Slutsats: H_0 förkastas på risknivån 1% eftersom $Z_{OBS} > Z_{0,005}$, $3,4831 > 2,5758$

R

Man kan alltså utläsa en skillnad mellan de två valdistrikten med avseende på andelen som tänkte rösta på den aktuella kandidaten.

R

*med avseende på andelen som tänkte rösta på den aktuella kandidaten

SU, STATISTIK

Skrivsal: VÅRTASALEN Anonymkod: SGK-0014 Blad nr: 4

4) 20p

UPPGIFT 4

	MÄN	KVINNOR	
N	50	75	125
L	25	100	125
G	75	50	125
H	50	75	125
	200	300	500 ← n

A

Ställ upp hypoteser. Eftersom man vill testa om det skiljer sig åt mellan män & kvinnor blir hypoteserna,

H_0: Inget beroende* H_1: Beroende*

R

B

Beslutsregel: H_0 förkastas om $\chi^2 > \chi^2_{\alpha[(c-1)(r-1)]}$

$\chi^2_{\alpha[(c-1)(r-1)]} = \chi^2_{0,01[(2-1)(4-1)]} = \chi^2_{0,01[3]}$

Från Tabell 4 ser vi att $\chi^2_{0,01(3)} = 11,345$ R

Vi använder Testvariabel:

$$\chi^2 = \sum \sum \frac{(n_{ij} - n_{i.} n_{.j}/n)^2}{n_{i.} n_{.j}/n}$$

V.G.V →

* mellan kön & vilken solprodukt som används

R

Vi summerar alltså ;

$$\frac{(50-125\times200/500)^2}{125\times200/500}=0$$

$$\frac{(75-125\times300/500)^2}{125\times300/500}=0$$

$$\frac{(25-125\times200/500)^2}{125\times200/500}=\frac{625}{50}=12,5$$

$$\frac{(100-125\times300/500)^2}{125\times300/500}=\frac{625}{75}\approx8,333$$

$$\frac{(75-125\times200/500)^2}{125\times200/500}=\frac{625}{50}=12,5$$

$$\frac{(50-125\times300/500)^2}{125\times300/500}=\frac{625}{75}\approx8,333$$

$$\frac{(50-125\times200/500)^2}{125\times200/500}=0$$

$$\frac{(75-125\times300/500)^2}{125\times300/500}=0$$

R

Summerar dessa:

$0+0+12,5+8,333+12,5+8,333+0+0=41,666$

Slutsats: H_0 förkastas på 1% signifikansnivå eftersom $\chi^2_{OBS} > \chi^2_{0,01[3]}$; $41,666 > 11,345$

R

Vi kan alltså säga på 1% signifikansnivå att det finns ett beroende mellan kön & vilken solprodukt som föredras.

R

5) 20p

UPPGIFT 5

	S1	S2
A1	15	20
A2	25	15
A3	8	13
A4	15	12
A5	30	5

utan sannolikheter till dom olika naturtillstånden (utfall) samma sannolikhet för samtliga tillstånd. I detta fall har vi två tillstånd alltså antas varje tillstånd inträffa med 50% sannolikhet.

A

Vilket alternativ väljs utifrån Laplace kriteriet

Laplacekriteriet menar att man väljer det alternativ med högst förväntat värde/nytta

Vi räknar förväntat värde för alla alternativ och väljer det med högst värde.

A1 : $(0{,}5 \times 15) + (0{,}5 \times 20) = 17{,}5$

$\boxed{\text{A2 : } (0{,}5 \times 25) + (0{,}5 \times 15) = 20}$

A3 : $(0{,}5 \times 8) + (0{,}5 \times 13) = 10{,}5$

A4 : $(0{,}5 \times 15) + (0{,}5 \times 12) = 13{,}5$

A5 : $(0{,}5 \times 30) + (0{,}5 \times 5) = 17{,}5$

Eftersom: $20 > 17{,}5 \geq 17{,}5 > 13{,}5 > 10{,}5$ väljs Alternativ 2 (A2)

R

R

V.G.V. →

B

Nu tilldelas S1 & S2 sannolikheter.

S1 → 0,8 & S2 → 0,2

Vi kan därför inte utgå från att det är lika mycket chans/risk för båda naturtillstånden som i uppgiften innan. Den nya uträkningen blir därför:

A1: $(0{,}8 \times 15) + (0{,}2 \times 20) = 16$

A2: $(0{,}8 \times 25) + (0{,}2 \times 15) = 23$

A3: $(0{,}8 \times 8) + (0{,}2 \times 13) = 9$

A4: $(0{,}8 \times 15) + (0{,}2 \times 12) = 14{,}4$

$\boxed{\text{A5: } (0{,}8 \times 30) + (0{,}2 \times 5) = 25}$

Eftersom $25 > 23 > 16 > 14{,}4 > 9$ väljs Alternativ 5 (A5).

(7)

Statistiska institutionen

Stockholms universitet

Rättningsblad

Datum: 5/6 - 2015

Sal: Värtasalen

Tenta: Statistikens grunder 2 (kväll)

Kurs: Statistikens grunder

ANONYMKOD: SGK-0003

☒ Jag godkänner att min tenta får läggas ut anonymt på hemsidan som studentsvar.

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 RC
Lär.ant. 20p	20p	20p	20p	20p					

POÄNG	BETYG	Lärarens sign.
100	A	RC

SU, STATISTIK

Skrivsal: Värtasalen Anonymkod: SGK-0003 Blad nr: 1

1) 20p

(1)

$X \sim N(\mu, \sigma)$, $n = 16$, $\bar{x} = 228$, $s = 5$, #frihetsgrader = 15

(a) $\bar{x} \pm t_{0,05}^{(15)} \frac{s}{\sqrt{n}} \approx 228 \pm 1,753 \frac{5}{4} = 228 \pm 2,19125$

Konfidensintervallet är $(225,8 \,;\, 230,2)$ R

R

(b) $H_0: \mu \geq 230$, $H_1: \mu < 230$

Testvariabel: $T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$

Förkasta H_0 om $t_{obs} < -t_{0,05}^{(15)} \approx -1,75$

R

$t_{obs} = \frac{228 - 230}{5/4} = -\frac{8}{5} = -1,6 > -1,75$

$\therefore$ H_0 kan ej förkastas. Det gick inte att påvisa att $\mu < 230$ på 5%-nivån.
(Slutsatsen kan även dras utifrån konfidensintervallet i (a).) R

(c) $p = P(X < 230) = P(T < -1,6) = P(T > 1,6)$

Värdet ligger mellan $\alpha = 0,10$ och $\alpha = 0,05$ med 15 frihetsgrader. R

$\therefore$ $0,05 < p < 0,1$

(2) 2) 2p

$n = 100, \ \hat{p} = 0{,}47$

(a) $\hat{p} \pm z_{0,025}\sqrt{\hat{p}(1-\hat{p})/n} \approx 0{,}47 \pm 1{,}96\sqrt{0{,}47 \cdot 0{,}53/100} \approx 0{,}47 \pm 0{,}098$

Konfidensintervallet är $(0{,}372; \ 0{,}568)$ R

(b) $H_0: p_0 \leq 0{,}40, \ H_1: p_0 > 0{,}40$

Testvariabel: $Z = \dfrac{\hat{p} - p_0}{\sqrt{p_0(1-p_0)/n}}$

Förkasta H_0 om $z_{obs} > z_{0,01} \approx 2{,}33$

$z_{obs} = \dfrac{0{,}47 - 0{,}40}{\sqrt{0{,}4 \cdot 0{,}6/100}} = \dfrac{0{,}07}{\sqrt{0{,}0024}} \approx 1{,}43 < 2{,}33$ R

R ∴ H_0 kan ej förkastas. Det gick inte att påvisa att $p > 0{,}40$ på 1%-nivån.

(Värdet 0,4 ingår även i konfidensintervallet, på 2,5%-nivå, ovan.) R

(c) p-värdet $= P(Z \geq 1{,}43) = 1 - P(Z < 1{,}43) = 1 - \Phi(1{,}43) \approx 1 - 0{,}9236 = 0{,}0764$

SU, STATISTIK

Skrivsal: Värtasalen Anonymkod: SGK-0003 Blad nr: 2

3) 20p

(3)

$n_1 = 120$, $Y_1 = 96$, $n_2 = 140$, $Y_2 = 84$, $\hat{p}_1 = \frac{96}{120} = \frac{4}{5}$, $\hat{p}_2 = \frac{84}{140} = \frac{3}{5}$, $(Y_1+Y_2)/(n_1+n_2) = \frac{180}{260} = \frac{9}{13}$

(a) Testvariabel: $Z = \dfrac{Y_1/n_1 - Y_2/n_2}{\sqrt{\frac{Y_1+Y_2}{n_1+n_2}\left(1-\frac{Y_1+Y_2}{n_1+n_2}\right)\left(\frac{1}{n_1}+\frac{1}{n_2}\right)}}$

$$z_{obs} = \frac{4/5 - 3/5}{\sqrt{\frac{9}{13}\left(1-\frac{9}{13}\right)\left(\frac{1}{120}+\frac{1}{140}\right)}} = \frac{1/5}{\sqrt{\frac{9}{13}\frac{4}{13}\frac{13}{840}}} = \frac{1/5}{\sqrt{\frac{3}{910}}} \approx 3{,}48 > z_{0{,}005} \approx 2{,}58$$ R

(b) H_0: $p_1 = p_2$, d.v.s. det finns ingen skillnad mellan de två valdistrikten

H_1: $p_1 \neq p_2$, d.v.s. det finns en skillnad mellan valdistrikten R

(c) H_0 kan förkastas på 1%-nivån eftersom $z_{obs} > z_{0{,}005}$

Vi kan med stor tillförlitlighet säga att det finns en skillnad mellan valdistrikten. R

4) 20p

(4)

(a) H_0: det är ingen skillnad mellan andelen män som föredrar resp produkt och andelen kvinnor som föredrar resp. produkt

R

H_1: det finns en skillnad (d.v.s. för några produkter skiljer sig andelen män som föredrar dem mot andelen kvinnor som gör det)

R

(b)

O_{ij}	M	K	Σ	p
N	50	75	125	0,25
O	25	100	125	0,25
G	75	50	125	0,25
H	50	75	125	0,25
Σ	200	300	500	1
p	0,4	0,6	1	

E_{ij}	M	K
N	50	75
O	50	75
G	50	75
H	50	75

frihetsgrader = (4-1)(2-1) = 3

O: Observerade frekvenser, E: Förväntade frekvenser

Testvariabel: $\chi^2 = \sum_i \sum_j \frac{(O_{ij}-E_{ij})^2}{E_{ij}}$

R

Förkasta H_0 om $\chi^2_{obs} > \chi^2_{0,01}(3) \approx 11,345$

$$\chi^2_{obs} = \frac{0+25^2+25^2+0}{50} + \frac{0+25^2+25^2+0}{75} = \frac{2\cdot 25^2}{50} + \frac{2\cdot 25^2}{75} = 25 + \frac{50}{3} = \frac{125}{3} \approx 41,7 > 11,345$$

∴ H_0 kan förkastas på 1%-nivån

R

Vi kan med stor tillförlitlighet säga att mäns och kvinnors preferenser fördelar sig olika mellan produkterna

R

SU, STATISTIK

Skrivsal: Vårtasalen Anonymkod: SGK-0003 Blad nr: 3

5) 20 p

(5)

	S_1	S_2	Förv. utilitet i (a)	Förv. utilitet i (b)
A_1	15	20	17,5	$0,8 \cdot 15 + 0,2 \cdot 20 = 16$
A_2	25	15	20	$0,8 \cdot 25 + 0,2 \cdot 15 = 23$
A_3	8	13	10,5	$0,8 \cdot 8 + 0,2 \cdot 13 = 9$
A_4	15	12	13,5	$0,8 \cdot 15 + 0,2 \cdot 12 = 14,4$
A_5	30	5	17,5	$0,8 \cdot 30 + 0,2 \cdot 5 = 25$
(a)	0,5	0,5		
(b)	0,8	0,2		

R

R

(a) Det högsta förväntade värdet under lika sannolikhet för naturtillstånden, 20, ges av A2.

(b) Det högsta förväntade värdet under de givna sannolikheterna, 25, ges av A5.

R