Stockholms
universitet
Statistiska institutionen
Raul Cano

SKRIVNINGSDATUM: 15-01-2020

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).
Återlämning av tentamen: hämtas på studentexpeditionen, plan 7 i B-huset fr.o.m. torsdagen den 23 januari. Kolla på vår hemsida studentexpeditionens mottagningstider under terminstid. Kontakta studentexpeditionen om du vill ha en skannad kopia av din rättade tentamen via e-post.

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)
Livslängden för en viss typ av elektriska komponenter kan antas variera som en normalfördelad stokastisk variabel med väntevärdet μ (timmar) och standardavvikelsen σ (timmar). För att undersöka värdet på μ har man från en mycket stor produktionsserie helt slumpmässigt valt ut 100 komponenter. Man har fått ett genomsnittsvärde $\bar{x} = 1040$ och en standardavvikelse s = 150.

a). Testa om μ är större än 1000. Genomför hypotesprövningen på signifikansnivå $\alpha = 0{,}05$. (5 p). Lös inte problemet via ett konfidensintervall, då lösningen ger 0 p.

b). Beräkna testets styrka (the power of the test) ifall μ är lika med 950. (10 p.)

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

Uppgift 2: (20 poäng)
I den stora staden New York och före ett presidentval uppgav 96 av 120 slumpmässigt utvalda röstberättigade personer 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). Beräkna ett 95% konfidensintervall för skillnaden mellan de två valdistrikten med avseende på andelen som skulle uppge ~~att de tänkte~~ rösta på kandidaten ifråga. (5 p.)

b). Testa på risknivån 1% ($\alpha = 0{,}01$) hypotesen att ovanstående skillnad är större än 0. Vilken blir din slutsats? (5 p.) Lös inte problemet via ett konfidensintervall, då lösningen ger 0 p.

c). Beräkna p-värdet i samband med ovanstående hypotesprövning. (10 p.)

Uppgift 3: (20 poäng)
I en marknadsundersökning studerar man skillnaden i konsumtion mellan tätort och landsbygd. För en viss förbrukningsartikel har man i de slumpmässigt valda hushållen under den senaste månaden fått följande resultat:

	Antal hushåll	Medel-värde	Standard-avvikelse
Tätort	250	190 kr	90 kr
Landsbygd	250	150 kr	30 kr

a). Beräkna ett 95% konfidensintervall för skillnaden i genomsnittlig konsumtion mellan tätortshushåll och landsbygdshushåll. (10 p.)

b). Testa på risknivån 1% ($\alpha = 0{,}01$) hypotesen att ovanstående skillnad är större än 0. Vilken blir din slutsats? (10 p.) Lös inte problemet via ett konfidensintervall, då lösningen ger 0 p.

Uppgift 4: (20 poäng)

Vid 192 kast med en tärning erhölls följande antal ettor, tvåor, etc: 30, 14, 18, 40, 52, 38.

a). Testa på 5% signifikansnivå om tärningen kan antas vara symmetrisk. (15 p.)

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

Uppgift 5: (20 poäng)

Låt oss anta följande beslutsmatris (enligt kapitel 29 i boken):

	Naturtillstånden		
Handlingsalternativ	S1	S2	S3
A1	6	8	22
A2	8	18	16
A3	12	12	12
A4	10	6	2

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

b). Anta att sannolikheterna för att de möjliga naturtillstånden inträffar är 0.3, 0.2 respektive 0.5 för S1, S2 respektive S3 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)

$E(c) = c$ $\quad$ $V(c) = 0$

$E(X + c) = E(X) + c$ $\quad$ $V(X + c) = V(X)$

$E(aX) = aE(X)$ $\quad$ $V(aX) = a^2V(X)$

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

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):

$$\bar{d} \pm t_{\alpha/2}^{(f.g.)} \frac{s_d}{\sqrt{n}}$$

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

$$\frac{y}{n} \pm z_{\alpha/2}\sqrt{(y/n)(1-y/n)/n}$$

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

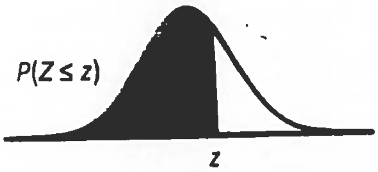
$$\hat{p}_1 - \hat{p}_2 \pm z_{\alpha/2}\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

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

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

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

TABELL 1. Normalfördelningen, standardiserad



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

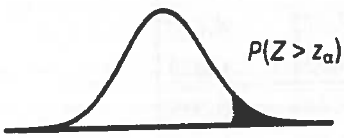
För negativa värden, utnyttja att $\Phi(-z) = 1 - \Phi(z)$.

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 $P(Z \leq -z_\alpha)$.

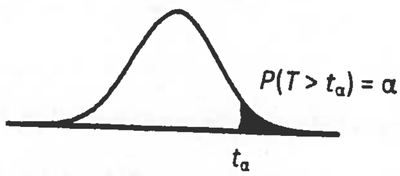


α	z_α
0,1	1,2816
0,05	1,6449
0,025	1,9600
0,010	2,3263
0,005	2,5758
0,0025	2,8070
0,0010	3,0902
0,0005	3,2905
0,00025	3,4808
0,00010	3,7190
0,00005	3,8906
0,000025	4,0556
0,000010	4,2649
0,000005	4,4172

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 $P(T \le -t_\alpha) = P(T > t_\alpha)$.



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.

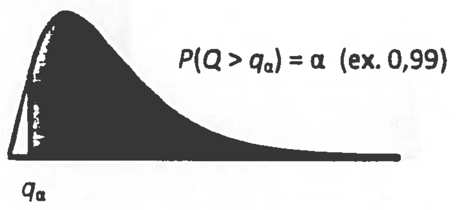
Vilket värde har q_α om $P(Q > q_\alpha) = \alpha$ där α är en sannolikhet.

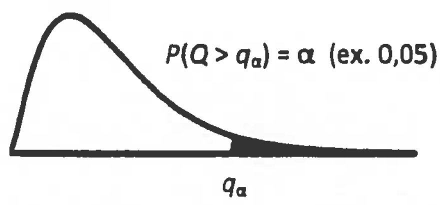
v	α = 0,999	0,995	0,99	0,975	0,95	0,05	0,025	0,01	0,005	0,001
1	0,000	0,000	0,000	0,001	0,004	3,841	5,024	6,635	7,879	10,828
2	0,002	0,010	0,020	0,051	0,103	5,991	7,378	9,210	10,597	13,816
3	0,024	0,072	0,115	0,216	0,352	7,815	9,348	11,345	12,838	16,266
4	0,091	0,207	0,297	0,484	0,711	9,488	11,143	13,277	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





X7

7

Stockholms universitet

Statistiska institutionen

Rättningsblad

Datum: 15/1-2020

Sal: Värtasalen

Tenta: Statistikens grunder 2, kväll

Kurs: Statistikens grunder, kväll

ANONYMKOD: 0004-HDY

☐ 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					

MW

POÄNG	BETYG	Lärarens sign.
100	A	RC

SU, STATISTIK

Skrivsal: VÅ Anonymkod: 0004-HDY Blad nr: 1

1) 20p

1 Låt X vara livslängden för typen av elektriska komponenter.

$X \sim N(\mu, \sigma^2)$

I en undersökning är $n = 100$, $\bar{x} = 1040$, $s = 150$

a) N är mkt stor och det behövs därför ingen ändlighetskorrektion då $\frac{N-n}{N-1} \approx 0$

$H_0: \mu \leq 1000$ $H_1: \mu > 1000$ $\alpha = 0{,}05$

Variansen ges av $\frac{s^2}{n} = \frac{150^2}{100} = 225$, $\frac{s}{\sqrt{n}} = 15$

R

Testvariabel är

$Z = \frac{\bar{X} - \mu}{s/\sqrt{n}}$ och eftersom mothypotesen är enkelsidig förkastas H_0 om

$Z_{obs} > Z_{0{,}025} = 1.6449$

$Z_{obs} = \frac{1040 - 1000}{15} = 2.667 > 1.6449$ R

H_0 förkastas därmed på en 5% sig. nivå till förmån av H_1. Det finns belägg för att μ är större än 1000 timmar. R

b) Testets styrka ges av $P(H_0 \text{ förkastas} \mid \mu = 950)$ om vi antar att $\mu = 950$ timmar.

$P(Z_{obs} > 1.6449 \mid \mu = 950) =$

→

$$P\left(\frac{\bar{X}-1000}{15} > 1.6449 \mid \mu = 950\right) = P(\bar{X} > 1024,7 \mid \mu = 950) =$$

$$P\left(Z > \frac{1024.7-950}{15}\right) = P(Z > 4.98) = 1 - P(Z < 4.98)$$

$\approx 0.$ R

Tabellen går endast till $z = 4$ och Sth att $z < 4.98$ är därför nästan 1, vilket innebär att testets styrka är ungefär 0. vid $\alpha = 0.05$. om $\mu = 950$

c) P-värdet ges av

$$P(Z > Z_{obs}) = P(Z > 2,667) \approx 1 - P(Z < 2,67)$$

$$= 1 - 0,99621 = 0,00379$$ R

P-värdet är ~~#~~ 0,38%.

P-värdet är det minsta värdet på α som H_0 skulle förkastas vid.

SU, STATISTIK

Skrivsal: VÅ Anonymkod: 0004-HDY Blad nr: 2

2) 2p

2. Låt P_A vara proportionen som tänker rösta på kandidaten i valdistrikt A, och P_B motsvarande siffra i distrikt B. Då $n(1-\hat{p})(\hat{p}) > 9$ för både A och B kan P_A och P_B skattas av $\hat{P}_A$ och $\hat{P}_B$ som antas ha följande fördelning:

$\hat{P}_A \sim N\left(P_A, \frac{P_A(1-P_A)}{n_A}\right)$

$\hat{P}_B \sim N\left(P_B, \frac{P_B(1-P_B)}{n_B}\right)$

$n_A = 120$ $\hat{P}_A = \frac{96}{120} = 0.8$ $\sigma_{\hat{P}_A} = \sqrt{\frac{0.8 \times 0.2}{120}} = 0.0365$

$n_B = 140$ $\hat{P}_B = \frac{84}{140} = 0.6$ $\sigma_{\hat{P}_B} = \sqrt{\frac{0.6 \times 0.4}{140}} = 0.0414$

Konfidensintervallet på 95% av $P_A - P_B$ ges av

K.I: $P_A - P_B \pm Z_{0,025} \sqrt{\sigma_{\hat{P}_A}^2 + \sigma_{\hat{P}_B}^2}$

$0,2 \pm 1,96\sqrt{0,0365^2 + 0,0414^2}$

$0,2 \pm 0.1082$ R

därmed är den sanna skillnaden $P_A - P_B$ med 95% sth inom intervallet

$[0,0918 \; ; \; 0,3082]$. R

b) $H_1: P_A > P_B \rightarrow P_A - P_B > 0$ $H_0: P_A = P_B \rightarrow P_A - P_B = 0$

testvariabeln ges av: $z = \frac{P_A - P_B}{\sqrt{P_p(1-P_p)\left(\frac{1}{n_A} + \frac{1}{n_B}\right)}}$

där P_p är den poolade proportionen och ges av

$P_p = \frac{84+96}{120+140} = 0,6923.$

Då hypotesen är enkelsidig ~~för~~ förkastas H_0 om

$Z_{obs} > Z_{0,01} = 2,3263.$

$$Z_{obs} = \frac{0,2}{\left((0.6923)(1-0.6923)\left(\frac{1}{120}+\frac{1}{140}\right)\right)^{\frac{1}{2}}} = 3,4833 > 2,3263$$

H_0 förkastas därmed till förmån av H_1 på sig. nivå 1%. Det finns belägg för att $P_A > P_B$.

c)

P-värdet av ovanstående hypotesprövning är det minsta värdet på α som H_0 skulle förkastas vid. Det ges av $P(Z > Z_{obs} \mid H_0 \text{ är sann}) =$

$P(Z > Z_{obs}) = P(Z > 3,4833) = 1 - P(Z < 3,4833) \approx$

$1 - P(Z < 3,48) = 1 - 0,99975 = 0,00025$

P-värdet är därför ~~0,00025%~~ 0,025%.

SU, STATISTIK

Skrivsal: VÅ **Anonymkod:** 0004-HDY **Blad nr:** 3

3) 20p

3. Låt μ_T (medelvärde i tätort) och μ_L (medelvärde Landsbygd).

μ_T och μ_L skattas av $\bar{X}_T$ och $\bar{X}_L$ då $n_T = n_L > 30$

Och $\bar{X}_T$ samt $\bar{X}_L$ kan antas ha följande fördelning.

$\bar{X}_T \sim N(\mu_T, \sigma_T^2)$ där σ_T^2 skattas av $\frac{S_T^2}{n_T}$

$\bar{X}_L \sim N(\mu_L, \sigma_L^2)$ där σ_L^2 skattas av $\frac{S_L^2}{n_L}$.

Vi är intresserade av $\mu_T - \mu_L$ som skattas av $\bar{X}_T - \bar{X}_L$. Ett konfidensintervall som med 95% s/h innehåller $\mu_T - \mu_L$ ges av:

$$\bar{X}_T - \bar{X}_L \pm Z_{0.025}\sqrt{\frac{S_T^2}{n_T} + \frac{S_L^2}{n_L}}, \quad Z_{0.025} = 1.96$$ R

$n_T = n_L = 250$ $\bar{x}_T = 190$ $\bar{x}_L = 150$ $S_T = 90$ $S_L = 30$

K.I: $190 - 150 \pm 1.96\sqrt{32.4 + 3.6}$

K.I: 40 ± 11.76 R

~~$\bar{X}_T$~~ $\mu_T - \mu_L$ ligger därför med 95% s/h ~~mellan~~ i intervallet $[28,24 ; 51,76]$. R

~~b) $\alpha = 0.01$~~

~~$H_0: |\mu_T - \mu_L| = 0$ $H_1: \mu_T - \mu_L \neq 0$~~

~~testvariabel:~~

~~$Z = \frac{\bar{x}_T - \bar{x}_L}{\sqrt{32.4 + 3.6}}$~~

~~Eftersom H_1 är dubbelsidig så förkastas H_0 i förmån till H_1 om $|Z_{obs}| > Z_{0,005} = 2.5758$~~

~~$Z_{obs} = \frac{190 - 150}{\sqrt{32.4 + 3.6}}$~~

$H_0: \mu_T - \mu_L = 0$ $H_1: \mu_T - \mu_L > 0$

$\alpha = 0,01$

testvariabel: $Z = \frac{\bar{x}_T - \bar{x}_L}{\sqrt{\frac{s_T^2}{n_T} + \frac{s_L^2}{n_L}}}$

H_1 är enkelsidig och H_0 förkastas därför om

$Z_{obs} > Z_{0,01} = 2,3263$ R

$Z_{obs} = \frac{190 - 150}{\sqrt{\frac{90^2}{250} + \frac{30^2}{250}}} = 6,667 > 2,3263$ R

H_0 förkastas till förmån av H_1. Det finns belägg för att medelsnittskonsumtionen i tätorter är större än landsbygden $(\mu_T > \mu_L)$ med risknivån 1%. R

SU, STATISTIK

Skrivsal: VÅ Anonymkod: 0004-HDY Blad nr: 4

4) 20p

4. ~~Antal~~

Sida (i)	1	2	3	4	5	6	totalt
Antal (n_i)	30	14	18	40	52	38	192
$E_i = n \cdot P_i$	32	32	32	32	32	32	192
P_i	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	1
$(E_i - n_i)^2$	2^2	18^2	14^2	8^2	20^2	6^2	1024

b) H_0: Tärningen är symmetrisk och $P_i = \frac{1}{6}$ för alla i.

H_1: Tärningen är ej symmetrisk och P_i är inte $\frac{1}{6}$ för alla i.

$\alpha = 0,05$

Om tärningen vore symmetrisk skulle alla sidor ha $P_i = \frac{1}{6}$ och det förväntade värdet för alla kolumner blir därmed $E_i = \frac{192}{6} = 32$ kast. Eftersom alla kategorier har $n > 5$ kan ett χ^2-test göras och χ^2-testvariabeln förväntas vara fördelad med $k-1 = 5$ frihetsgrader.

Testvariabel $\chi^2 = \sum_{i=1}^{6} \frac{(n_i - E_i)^2}{E_i}$

H_0 förkastas om $\chi^2 > \chi^2_{(0,05)}(5) = 11,070$

$$\chi^2 = \frac{2^2 + 18^2 + \ldots + 6^2}{32} = \frac{1024}{32} = 32 > 11,070$$

Eftersom $Z_{obs} > 11.070$ förkastas H_0 och det finns belägg på sig.nivå 5% att tärningen inte är symmetrisk.

SU, STATISTIK

Skrivsal: VÅ Anonymkod: 0004 - HDY Blad nr: 5

5) 20p

5.

Handlingsalternativ	Naturtillstånd S_1	S_2	S_3	totalt
A_1	6	8	22	36
A_2	8	18	16	42
A_3	12	12	12	36
A_4	10	6	2	18

a) Laplacekriteriet innebär att man antar att slh att ~~[illegible]~~ naturtillstånden inträffar är samma för samtliga tillstånd. $P(S_1) = P(S_2) = P(S_3) = \frac{1}{3}$.

Jag antar att ~~en~~ det som ska uppnås är en hög siffra (det framgår inte i uppgiften vad siffrorna betyder), det vill säga att ett resultat på tex 8 är bättre än på 2. För att bestämma vilket handelsalternativ som bör väljas räknas därför det förväntade värdet av varje alternativ ut.

$E(A_1) = \frac{1}{3}(6+8+22) = 12$

$E(A_2) = \frac{1}{3}(42) = 14$ ← därmed bör A_2 väljas enl. Laplacekriteriet om högt resultat är önskvärt.

$E(A_3) = \frac{1}{3}(36) = 12$

$E(A_4) = \frac{1}{3}(18) = 6$

R

→

b) $P(S_1) = 0,3$

$P(S_2) = 0,2$

$P(S_3) = 0,5$

För att besluta om det bästa alternativet räknas det förväntade värdet för varje alternativ.

$E(x) = P_1(n_1) + P_2(n_2) + P_3(n_3) + \ldots$

$E(A_1) = 0.3 \times 6 + 0.2 \times 8 + 0,5 \times 22 = 14,4$ ← R

$E(A_2) = 0,3 \times 8 + 0,2 \times 18 + 0,5 \times 16 = 14$ R

$E(A_3) = 0,3 \times 12 + 0.2 \times 12 + 0,5 \times 12 = 12$

$E(A_4) = 0,3 \times 10 + 0,2 \times 6 + 0,5 \times 2 = 5,2$

$E(A_1)$ är störst givet att $P(S_1) = 0,3$, $P(S_2) = 0,2$ och $P(S_3) = 0,5$. Därför bör alternativ nummer 1 väljas.

R