Tentamentsskrivning i Statistisk slutledning MVE155/MSG200, 7.5 hp.

Tid: 7 juni 2017, kl 14.00-18.00
Examinator och jour: Serik Sagitov, tel. 031-772-5351, rum H3026 i MV-huset.
Hjälpmedel: Chalmersgodkänd räknare, **egen** formelsamling (fyra A4 sidor).
CTH: för "3" fordras 12 poäng, för "4" - 18 poäng, för "5" - 24 poäng.
GU: för "G" fordras 12 poäng, för "VG" - 20 poäng.
Inclusive eventuella bonuspoäng.

Partial answers and solutions are also welcome. Good luck!

1. (5 points) A stratified random sample of 700 households in a community was selected to estimate the mean μ of home improvement expenditures (in US dollars) last year. The strata sizes and sample results were as follows.

Stratum	Stratums size	Sample size	Sample mean	Sample st. deviation
1	12000	300	480	200
2	6000	150	380	150
3	10000	250	510	180

(a) The strata employed were geographic areas. Why does such a stratification design make sense?

(b) Give a point estimate of μ and compute its standard error.

(c) The difference between the largest and the smallest sample means is 130. Is it statistically significant?

2. (5 points) A multiple regression model

$$E(Y) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3$$

was used to explain the arthritis incidence measured by the the number of cases per 1000 inhabitants. Three explanatory variables were
x_1 = percent of population over 65 years old,
x_2 = number of physicians (per 1000),
x_3 = mean disposable income for the people over 65 (in thousands of US dollars).
A computer output for data collected in 13 different districts, has given four least squares estimates and their standard errors:

Parameter	Point estimate	Standard error
β_0	0.43914	1.57976
β_1	0.46963	0.11035
β_2	1.49976	0.67926
β_3	0.05921	0.08163

The sum of squares were computed to be: total = 28.51451, residuals = 4.41567, regression = 23.99884.

(a) Find 95% confidence intervals for β_2 and β_3. What are your conclusions about the choice of explanatory variables?

(b) For the simple regression model with a single explanatory variable x_1, the same data gave

Parameter	Point estimate	Standard error
β_0	2.38250	1.30464
β_1	0.56714	0.09719

The sum of squares for the residuals = 6.96186. Compare the multiple regression model to the simple regression model using the coefficient of determination.

(c) Compute a prediction interval for the arthritis incidence in a district with 20 percent of people being older than 65.

3. (5 points) Consider the following random sample of size 6

$$x_1 = 1.42, \quad x_2 = 0.58, \quad x_3 = -0.36, \quad x_4 = 3.76, \quad x_5 = 2.36, \quad x_6 = -1.76.$$

(a) Find a point estimate for the population median.

(b) Explain how would you estimate the standard error of your point estimate using a computer.

(c) The interval $(-1.76, 3.76)$ can be viewed as a 94% confidence interval for the median. Explain.

4. (5 points) Respond to the following.

(a) I have two batches of numbers and I have a corresponding $\bar{x}$ and $\bar{y}$. Why should I test whether they are equal when I can just see whether they are or not?

(b) Using the t-distribution is ridiculous. It is valid when the populations are normal and have equal variance. If the samples were so small that the t-distribution were practically different from the normal distribution, you would be unable to check these assumptions.

(c) A study of nearly 4000 elderly North Carolinians has found that those who attended religious services every week were 46% less likely to die over a six-year period than people who attended less often or not at all, according to researchers at Duke University Medical Center.

(d) The following table shows admission rates for the six most popular majors on the Berkeley campus.

Major	Male applicants	% admitted	Female applicants	% admitted
A	825	62	108	82
B	560	63	25	68
C	325	37	593	34
D	417	33	375	35
E	191	28	393	34
F	373	6	341	7

If the percentages admitted are compared, women do not seem to be unfavorably treated. But when the combined admission rated for all six majors are calculated, it is found that 44% of the men and only 30% of the women were admitted.

5. (5 points) Let the unknown probability that a basketball player makes a shot successfully be p. Suppose your prior on p is uniform on [0, 1] and that she then makes three successful shots in a row.

(a) What is the posterior density of p? What are your assumptions?

(b) What would you estimate the probability that she makes a fourth shot to be?

(c) What is a clear advantage of using the Bayesian approach for this type of settings?

6. (5 points) The table below gives the energy use of five gas cooking ovens for seven menu days. The units are equivalent kilowatt-hours: 0.24 kwh =1 ft^3 of natural gas.

Menu day	Oven 1	Oven 2	Oven 3	Oven 4	Oven 5
1	8.25	8.26	6.55	8.21	6.69
2	5.12	4.81	3.87	4.81	3.99
3	5.32	4.37	3.76	4.67	4.37
4	8.00	6.50	5.38	6.51	5.60
5	6.97	6.26	5.03	6.40	5.60
6	7.65	5.84	5.23	6.24	5.73
7	7.86	7.31	5.87	6.64	6.03

(a) Describe in detail an appropriate normal theory additive model. What are the main effects? Mention a number of possible noise factors.

(b) Clearly state two null hypotheses relevant for this data set, as well as their alternatives. Which one of the null hypotheses is more of interest? Explain.

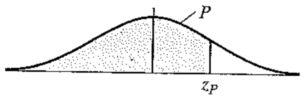
(c) Explain how a normal probability plot should be constructed to verify the normality assumption of the model.

(d) A relevant non-parametric test applied to this data using Matlab resulted in the P-value = 4.3695e-05, and something called

meanranks: [4.8571 3.4286 1 3.6429 2.0714].

Explain how the test works and the role of the vector meanranks.

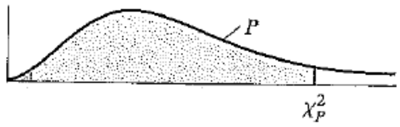
TABLE 2 Cumulative Normal Distribution—Values of P Corresponding to z_p for the Normal Curve



z is the standard normal variable. The value of P for $-z_p$ equals 1 minus the value of P for $+z_p$; for example, the P for -1.62 equals $1-.9474 = .0526$.

z_p	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
.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	.9772	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.1	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.2	.9861	.9864	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.3	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.4	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.5	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.6	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.7	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.8	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9979	.9980	.9981
2.9	.9981	.9982	.9982	.9983	.9984	.9984	.9985	.9985	.9986	.9986
3.0	.9987	.9987	.9987	.9988	.9988	.9989	.9989	.9989	.9990	.9990
3.1	.9990	.9991	.9991	.9991	.9992	.9992	.9992	.9992	.9993	.9993
3.2	.9993	.9993	.9994	.9994	.9994	.9994	.9994	.9995	.9995	.9995
3.3	.9995	.9995	.9995	.9996	.9996	.9996	.9996	.9996	.9996	.9997
3.4	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9997	.9998

TABLE 3 Percentiles of the χ^2 Distribution—Values of χ^2_P Corresponding to P



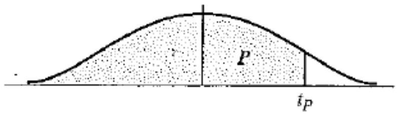
df	$\chi^2_{.005}$	$\chi^2_{.01}$	$\chi^2_{.025}$	$\chi^2_{.05}$	$\chi^2_{.10}$	$\chi^2_{.90}$	$\chi^2_{.95}$	$\chi^2_{.975}$	$\chi^2_{.99}$	$\chi^2_{.995}$
1	.000039	.00016	.00098	.0039	.0158	2.71	3.84	5.02	6.63	7.88
2	.0100	.0201	.0506	.1026	.2107	4.61	5.99	7.38	9.21	10.60
3	.0717	.115	.216	.352	.584	6.25	7.81	9.35	11.34	12.84
4	.207	.297	.484	.711	1.064	7.78	9.49	11.14	13.28	14.86
5	.412	.554	.831	1.15	1.61	9.24	11.07	12.83	15.09	16.75
6	.676	.872	1.24	1.64	2.20	10.64	12.59	14.45	16.81	18.55
7	.989	1.24	1.69	2.17	2.83	12.02	14.07	16.01	18.48	20.28
8	1.34	1.65	2.18	2.73	3.49	13.36	15.51	17.53	20.09	21.96
9	1.73	2.09	2.70	3.33	4.17	14.68	16.92	19.02	21.67	23.59
10	2.16	2.56	3.25	3.94	4.87	15.99	18.31	20.48	23.21	25.19
11	2.60	3.05	3.82	4.57	5.58	17.28	19.68	21.92	24.73	26.76
12	3.07	3.57	4.40	5.23	6.30	18.55	21.03	23.34	26.22	28.30
13	3.57	4.11	5.01	5.89	7.04	19.81	22.36	24.74	27.69	29.82
14	4.07	4.66	5.63	6.57	7.79	21.06	23.68	26.12	29.14	31.32
15	4.60	5.23	6.26	7.26	8.55	22.31	25.00	27.49	30.58	32.80
16	5.14	5.81	6.91	7.96	9.31	23.54	26.30	28.85	32.00	34.27
18	6.26	7.01	8.23	9.39	10.86	25.99	28.87	31.53	34.81	37.16
20	7.43	8.26	9.59	10.85	12.44	28.41	31.41	34.17	37.57	40.00
24	9.89	10.86	12.40	13.85	15.66	33.20	36.42	39.36	42.98	45.56
30	13.79	14.95	16.79	18.49	20.60	40.26	43.77	46.98	50.89	53.67
40	20.71	22.16	24.43	26.51	29.05	51.81	55.76	59.34	63.69	66.77
60	35.53	37.48	40.48	43.19	46.46	74.40	79.08	83.30	88.38	91.95
120	83.85	86.92	91.58	95.70	100.62	140.23	146.57	152.21	158.95	163.64

For large degrees of freedom,

$$\chi^2_P = \tfrac{1}{2}(z_P + \sqrt{2v-1})^2 \text{ approximately,}$$

where v = degrees of freedom and z_P is given in Table 2.

TABLE 4 Percentiles of the t Distribution



df	$t_{.60}$	$t_{.70}$	$t_{.80}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$
1	.325	.727	1.376	3.078	6.314	12.706	31.821	63.657
2	.289	.617	1.061	1.886	2.920	4.303	6.965	9.925
3	.277	.584	.978	1.638	2.353	3.182	4.541	5.841
4	.271	.569	.941	1.533	2.132	2.776	3.747	4.604
5	.267	.559	.920	1.476	2.015	2.571	3.365	4.032
6	.265	.553	.906	1.440	1.943	2.447	3.143	3.707
7	.263	.549	.896	1.415	1.895	2.365	2.998	3.499
8	.262	.546	.889	1.397	1.860	2.306	2.896	3.355
9	.261	.543	.883	1.383	1.833	2.262	2.821	3.250
10	.260	.542	.879	1.372	1.812	2.228	2.764	3.169
11	.260	.540	.876	1.363	1.796	2.201	2.718	3.106
12	.259	.539	.873	1.356	1.782	2.179	2.681	3.055
13	.259	.538	.870	1.350	1.771	2.160	2.650	3.012
14	.258	.537	.868	1.345	1.761	2.145	2.624	2.977
15	.258	.536	.866	1.341	1.753	2.131	2.602	2.947
16	.258	.535	.865	1.337	1.746	2.120	2.583	2.921
17	.257	.534	.863	1.333	1.740	2.110	2.567	2.898
18	.257	.534	.862	1.330	1.734	2.101	2.552	2.878
19	.257	.533	.861	1.328	1.729	2.093	2.539	2.861
20	.257	.533	.860	1.325	1.725	2.086	2.528	2.845
21	.257	.532	.859	1.323	1.721	2.080	2.518	2.831
22	.256	.532	.858	1.321	1.717	2.074	2.508	2.819
23	.256	.532	.858	1.319	1.714	2.069	2.500	2.807
24	.256	.531	.857	1.318	1.711	2.064	2.492	2.797
25	.256	.531	.856	1.316	1.708	2.060	2.485	2.787
26	.256	.531	.856	1.315	1.706	2.056	2.479	2.779
27	.256	.531	.855	1.314	1.703	2.052	2.473	2.771
28	.256	.530	.855	1.313	1.701	2.048	2.467	2.763
29	.256	.530	.854	1.311	1.699	2.045	2.462	2.756
30	.256	.530	.854	1.310	1.697	2.042	2.457	2.750
40	.255	.529	.851	1.303	1.684	2.021	2.423	2.704
60	.254	.527	.848	1.296	1.671	2.000	2.390	2.660
120	.254	.526	.845	1.289	1.658	1.980	2.358	2.617
∞	.253	.524	.842	1.282	1.645	1.960	2.326	2.576