14.30 Introduction to Statistical Methods in Economics
Spring 2009

14.30 Exam 3

Spring 2008

Instructions: This exam is closed-book and closed-notes. You may use a calculator and a cheat sheet. Please read through the exam first in order to ask clarifying questions and to allocate your time appropriately. In order to receive partial credit in the case of computational errors, please show all work. You have 90 minutes to complete the exam. Good luck!

1. (25 points) Expectations and Variances

Let

$$X \sim N(\mu_X, \sigma_X^2), \text{ and } Y \sim N(\mu_Y, \sigma_Y^2)$$

be normally distributed random variables with $\text{Cov}(X, Y) = \sigma_{XY}$. We look at a weighted sum of the two random variables, $Z = wX + (1 - w)Y$ for some constant w.

(a) As a function of w, what is the expectation of Z?

(b) As a function of w, what is the variance of Z?

(c) Calculate $\mathbb{E}[Z^2]$.

(d) Suppose you want to invest your savings, and X and Y are the returns to stock 1 and 2 respectively. If the fraction of your savings invested in stock 1 is w, your total return will be $Z = wX + (1 - w)Y$. Suppose both assets have the same *expected return* $\mu_X = \mu_Y$, and you'd like to chose w to keep the variation of your return as small as possible. Given your answer to (b), which value of w minimizes the variance of Z? Suppose X and Y are independent. When is the optimal w not between zero and one? Why shouldn't you just invest everything in the stock with the lower variance?

(e) (optional, extra credit): Suppose that $X = -\infty$ with probability p_X, and $X \sim N(\mu_X, \sigma_X^2)$ with probability $1 - p_X$. Similarly, $Y = -\infty$ with probability p_Y, and $Y \sim N(\mu_Y, \sigma_Y^2)$ with probability $1 - p_Y$. Also, X and Y are independent as before. Again, you have a combination of assets $wX + (1 - w)Y$, and it seems reasonable that you'd want to keep the probability $P(wX + (1 - w)Y = -\infty)$ as small as possible. What is the probability of such a shock given w? Would you still want to diversify? Does this contradict what you found in part (d)? A short answer suffices, no algebra needed.

2. (25 points) Estimation

You and your friend are waiting for your printing jobs in the Athena cluster. There are many other students in the lab, and the number of printing jobs in a 10-minute interval follows a Poisson distribution with (yet unknown) arrival rate λ.

Suppose that each time you run into your friend, you have already picked up your print-outs and are just having a 10 minute long chat with your friend. You keep track of the number X_i of printouts coming out of the printer during each of those 10-minute time intervals. For a sample $X_1, \ldots, X_{12}$, you find that the mean is $\bar{X}_{12} = 5.17$, and the mean of squares is $\frac{1}{n}\sum_{i=1}^{n} X_i^2 = 31.43$.

(a) Calculate the method of moments estimator for the Poisson rate λ using the sample mean.

(b) Recall that for a Poisson random variable X, $\text{Var}(X) = \lambda$. Can you construct an alternative method of moments estimator based on the information given above? Is that estimator different from that in (a)? If yes, do you think that this is a problem?

Athena is MIT's UNIX-based computing environment. OCW does not provide access to it.

Now, instead of counting the number of printing jobs, you only keep track of the time spent waiting until your friend's printing job came out. Your statistic textbook tells you that the waiting time T_i for the rth occurrence of a Poisson event has the p.d.f.

$$f_T(t) = \begin{cases} \frac{\lambda^r}{(r-1)!} t^{r-1} e^{-\lambda t} & \text{if } t > 0 \\ 0 & \text{otherwise} \end{cases}$$

You know that your friend's printing job is always 5 positions behind your own in the queue (i.e. $r = 5$) and you only keep track of the time T_i you are waiting together.

(c) What is the likelihood and the log-likelihood function for the i.i.d. sample $T_1, \dots, T_n$?

(d) Find the maximum likelihood estimator for λ.

(e) You lost your detailed notes on the sample $T_1, \dots, T_n$, but you still remember that the average observed waiting time was $\bar{T}_n = 3.84$, and the sample variance of waiting times, $\hat{S}_n = 4.12$. Will you still be able to compute a consistent estimator for λ?

3. (15 points) Confidence Intervals

Suppose you have an i.i.d. sample $X_1, \dots, X_{25}$ of 25 observations, where $X_i \sim N(\mu, \sigma^2)$. The mean of the sample is $\bar{X}_{25} = 1.2$.

(a) Construct a 95% confidence interval for μ, assuming that $\sigma^2 = 4$.

(b) Now suppose we didn't know σ^2, but I told you that $\bar{X^2}_{25} = \frac{1}{25} \sum_{i=1}^{25} X_i^2 = 6.38$. Construct a 95% confidence interval for μ.

(c) Suppose you want to test the null hypothesis $H_0 : \mu = 0$ against $H_A : \mu > 0$ at the 5% significance level. Given $\bar{X}_{25}$ and $\overline{X^2_{25}}$ as before (and assuming that you don't know the true variance σ^2), do you reject the null hypothesis?

4. (20 points) Hypothesis Tests

Suppose $X \sim U[-\theta, \theta]$. You have only one single observation from that distribution, and you want to test

$$H_0 : \theta = 1 \text{ against } H_A : \theta = 0.1$$

(a) State the p.d.f. of X (hint: be careful about where the p.d.f. is zero and where it's not).

(b) Derive the most powerful test - which result from the class are you using to show that the test is indeed most powerful?

(c) Consider the test "reject H_0 if $|X| < 0.1$". What is the size α of this test? What is its power $1 - \beta$? Be sure to give the definition of size and power when you do the calculation.

(d) Can you construct a most powerful test of size 5%?

5. (15 points) Hypothesis Tests

You are conducting a randomized experiment on telepathy in which your experimental subjects have to take a test. A medium, i.e. a person with supernatural abilities, claims that she is able to influence other peoples' thoughts and feelings only through extra-sensory perception.
You split your N experimental subjects into two groups of m and n individuals, respectively. For the first m individuals ("treatment group"), the medium is sitting in a room next door with the solutions manual to the test, thinking about the test-taker in order to help him/her on the test. The second group of n individuals ("controls") takes the test under normal conditions. In order not to boost the confidence of either group, participants in neither group are told whether they are helped by the medium.
You observe the samples $X_1, \ldots, X_m$ of scores for the group which received help from the medium, and $Y_1, \ldots, Y_n$ are the scores of the control group. You know beforehand that

$$X_i \sim N(\mu_X, \sigma^2), \text{ and } Y_i \sim N(\mu_Y, \sigma^2)$$

Each individual is tested separately and under identical conditions, so we assume that all observations are independent.

(a) Give the variances of $\bar{X}_m$ and $\bar{Y}_n$. What is the variance of $\bar{X}_m - \bar{Y}_n$?

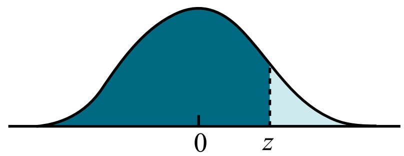
(b) For a given *total* number of experimental subjects $N = m + n$, which value of m minimizes the variance of the difference $\bar{X}_m - \bar{Y}_n$?

(c) You want to conduct a test of $H_0 : \mu_X = \mu_Y$ against $H_A : \mu_X \neq \mu_Y$ with size $\alpha = 5\%$. Suppose $\sigma^2 = 4$, $m = 20$ and $n = 8$, and you find that $\bar{X}_m - \bar{Y}_n = 1.74$. Do you reject the null hypothesis?

(d) Using your results from above, how many additional participants would you have to add at least in order to distinguish an effect of $\mu_X - \mu_Y = 1$ points from $\mu_X - \mu_Y = 0$ points with probability 95%? Hint: you can add subjects to the treatment group, the control group, or both (whichever requires the smallest number of additional participants).

Have a great Summer!

Cumulative areas under the standard normal distribution



(Cont.)

z	0	1	2	3	4	5	6	7	8	9
-3	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
-2.9	0.0019	0.0018	0.0017	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
-2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
-2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
-2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
-2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
-2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
-2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
-2.2	0.0139	0.0136	0.0132	0.0129	0.0126	0.0122	0.0119	0.0116	0.0113	0.0110
-2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
-2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
-1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0238	0.0233
-1.8	0.0359	0.0352	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0300	0.0294
-1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
-1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
-1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0570	0.0559
-1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0722	0.0708	0.0694	0.0681
-1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823
-1.2	0.1151	0.1131	0.1112	0.1093	0.1075	0.1056	0.1038	0.1020	0.1003	0.0985
-1.1	0.1357	0.1335	0.1314	0.1292	0.1271	0.1251	0.1230	0.1210	0.1190	0.1170
-1.0	0.1587	0.1562	0.1539	0.1515	0.1492	0.1469	0.1446	0.1423	0.1401	0.1379
-0.9	0.1841	0.1814	0.1788	0.1762	0.1736	0.1711	0.1685	0.1660	0.1635	0.1611
-0.8	0.2119	0.2090	0.2061	0.2033	0.2005	0.1977	0.1949	0.1922	0.1894	0.1867
-0.7	0.2420	0.2389	0.2358	0.2327	0.2297	0.2266	0.2236	0.2206	0.2177	0.2148
-0.6	0.2743	0.2709	0.2676	0.2643	0.2611	0.2578	0.2546	0.2514	0.2483	0.2451
-0.5	0.3085	0.3050	0.3015	0.2981	0.2946	0.2912	0.2877	0.2843	0.2810	0.2776
-0.4	0.3446	0.3409	0.3372	0.3336	0.3300	0.3264	0.3228	0.3192	0.3156	0.3112
-0.3	0.3821	0.3783	0.3745	0.3707	0.3669	0.3632	0.3594	0.3557	0.3520	0.3483
-0.2	0.4207	0.4168	0.4129	0.4090	0.4052	0.4013	0.3974	0.3936	0.3897	0.3859
-0.1	0.4602	0.4562	0.4522	0.4483	0.4443	0.4404	0.4364	0.4325	0.4286	0.4247
-0.0	0.5000	0.4960	0.4920	0.4880	0.4840	0.4801	0.4761	0.4721	0.4681	0.4641

Image by MIT OpenCourseWare.

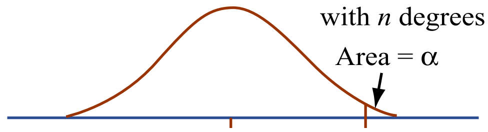
Cumulative areas under the standard normal distribution

(Cont.)

z	0	1	2	3	4	5	6	7	8	9
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7703	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9278	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9430	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9648	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9700	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9762	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9874	0.9878	0.9881	0.9884	0.9887	0.9890
2..3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	09988	0.9989	0.9989	0.9989	0.9990	0.9990

Source: B. W. Lindgren, *Statistical Theory* (New York: Macmillan. 1962), pp. 392-393.

Upper Percentiles of Student t Distributions



	α						
df	0.20	0.15	0.10	0.05	0.025	0.01	0.005
1	1.376	1.963	3.078	6.3138	12.706	31.821	63.657
2	1.061	1.386	1.886	2.9200	4.3027	6.965	9.9248
3	0.978	1.250	1.638	2.3534	3.1825	4.541	5.8409
4	0.941	1.190	1.533	2.1318	2.7764	3.747	4.6041
5	0.920	1.156	1.476	2.0150	2.5706	3.365	4.0321
6	0.906	1.134	1.440	1.9432	2.4469	3.143	3.7074
7	0.896	1.119	1.415	1.8946	2.3646	2.998	3.4995
8	0.889	1.108	1.397	1.8595	2.3060	2.896	3.3554
9	0.883	1.100	1.383	1.8331	2.2622	2.821	3.2498
10	0.879	1.093	1.372	1.8125	2.2281	2.764	3.1693
11	0.876	1.088	1.363	1.7959	2.2010	2.718	3.1058
12	0.873	1.083	1.356	1.7823	2.1788	2.681	3.0545
13	0.870	1.079	1.350	1.7709	2.1604	2.650	3.0123
14	0.868	1.076	1.345	1.7613	2.1448	2.624	2.9768
15	0.866	1.074	1.341	1.7530	2.1315	2.602	2.9467
16	0.865	1.071	1.337	1.7459	2.1199	2.583	2.9208
17	0.863	1.069	1.333	1.7396	2.1098	2.567	2.8982
18	0.862	1.067	1.330	1.7341	2.1009	2.552	2.8784
19	0.861	1.066	1.328	1.7291	2.0930	2.539	2.8609
20	0.860	1.064	1.325	1.7247	2.0860	2.528	2.8453
21	0.859	1.063	1.323	1.7207	2.0796	2.518	2.8314
22	0.858	1.061	1.321	1.7171	2.0739	2.508	2.8188
23	0.858	1.060	1.319	1.7139	2.0687	2.500	2.8073
24	0.857	1.059	1.318	1.7109	2.0639	2.492	2.7969
25	0.856	1.058	1.316	1.7081	2.0595	2.485	2.7874
26	0.856	1.058	1.315	1.7056	2.0555	2.479	2.7787
27	0.855	1.057	1.314	1.7033	2.0518	2.473	2.7707
28	0.855	1.056	1.313	1.7011	2.0484	2.467	2.7633
29	0.854	1.055	1.311	1.6991	2.0452	2.462	2.7564
30	0.854	1.055	1.310	1.6973	2.0423	2.457	2.7500
31	0.8535	1.0541	1.3095	1.6955	2.0395	2.453	2.7441
32	0.8531	1.0536	1.3086	1.6939	2.0370	2.449	2.7385
33	0.8527	1.0531	1.3078	1.6924	2.0345	2.445	2.7333
34	0.8524	1.0526	1.3070	1.6909	2.0323	2.441	2.7284

Image by MIT OpenCourseWare.