X_400004 - Statistics
Solutions to the Midterm Resit

9 February 2021

Below are answers to the exam questions. Some of these are slightly abbreviated, while others include extra comments. These are for your reference only but should inform the level of detail that is expected from your answers in the exam. Also keep in mind that there might be different ways to approach each question. If you find typos and or omissions, please report them to the lecturer so they can be corrected.

Prob.I: In the study of wireless communication networks one often comes across a process that is well modeled by samples from continuous random variables with density

$$f(x) = \begin{cases} \frac{1}{\sqrt{\lambda}} e^{-\frac{x}{\sqrt{\lambda}}} & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases},$$

where $\lambda > 0$ is a parameter we would like to estimate. Suppose you observe two **independent** samples X_1 and X_2 from the above distribution and consider the following two possible estimators of λ:

$$\hat{\lambda} = \frac{X_1^2 + X_2^2}{4} \qquad \text{and} \qquad \tilde{\lambda} = \frac{X_1^2 + X_2^2 + 2X_1X_2}{6}.$$

You can use the fact that, if X is a random variable with the density above then

$$\mathbb{E}(X) = \sqrt{\lambda}, \qquad \mathbb{E}(X^2) = 2\lambda, \qquad \mathbb{E}(X^3) = 6\lambda\sqrt{\lambda}, \qquad \mathbb{E}(X^4) = 24\lambda^2.$$

14 pts (a) Compute the bias of the two estimators. Is there an estimator that is biased?

Solution: First compute

$$\mathbb{E}(\hat{\lambda}) = \mathbb{E}\left(\frac{X_1^2 + X_2^2}{4}\right) = \frac{\mathbb{E}(X_1^2) + \mathbb{E}(X_2^2)}{4} = \frac{2\lambda + 2\lambda}{4} = \lambda.$$

The bias of this estimator is

$$\mathbf{bias}_{\hat{\lambda}}(\lambda) = \lambda - \lambda = 0.$$

To compute the bias of the second estimator note that X_1 and X_2 are independent, therefore

$$\mathbb{E}(\tilde{\lambda}) = \mathbb{E}\left(\frac{X_1^2 + X_2^2 + 2X_1X_2}{6}\right) = \frac{\mathbb{E}(X_1^2) + \mathbb{E}(X_2^2) + 2\mathbb{E}(X_1)\mathbb{E}(X_2)}{6} = \frac{2\lambda + 2\lambda + 2\lambda}{6} = \lambda.$$

Therefore

$$\mathbf{bias}_{\tilde{\lambda}}(\lambda) = \lambda - \lambda = 0.$$

Neither estimator is biased.

12 pts (b) Compute the variance of $\hat{\lambda}$.

Solution:

$$\mathbb{V}(\hat{\lambda}) = \mathbb{V}\left(\frac{X_1^2 + X_2^2}{4}\right) \overset{\text{ind.}}{=} \frac{\mathbb{V}(X_1^2) + \mathbb{V}(X_2^2)}{4^2} = \frac{1}{8}\mathbb{V}(X_1^2).$$

Now $\mathbb{V}(X_1^2) = \mathbb{E}(X_1^4) - \left(\mathbb{E}(X_1^2)\right)^2 = 20\lambda^2$**. Therefore** $\mathbb{V}(\hat{\lambda}) = \frac{5}{2}\lambda^2$**.**

6 pts (c) Compute the MSE of estimator $\hat{\lambda}$.

Solution: The MSE can be computed by summing the bias squared and the variance. Therefore

$$\textbf{MSE}_{\hat{\lambda}}(\lambda) = (0)^2 + \frac{5}{2}\lambda^2 = \frac{5}{2}\lambda^2.$$

6 pts (d) It can be shown that the MSE of $\tilde{\lambda}$ is given by $\frac{7}{3}\lambda^2$. Given this and your answer to (c) which of the two estimators would you prefer? Justify your answer.

Solution: Since $\textbf{MSE}_{\hat{\lambda}}(\lambda) = \frac{5}{2}\lambda^2 > \frac{7}{3}\lambda^2 = \textbf{MSE}_{\tilde{\lambda}}(\lambda)$ **the second estimator is preferable if the performance metric of interest is the mean squared error.**

Prob.II: Suppose that X is a discrete random variable that with the following distribution

$$\mathbb{P}(X=0)=\frac{2}{3}\theta, \quad \mathbb{P}(X=1)=\frac{1}{3}\theta, \quad \mathbb{P}(X=2)=\frac{2}{3}(1-\theta), \quad \mathbb{P}(X=3)=\frac{1}{3}(1-\theta),$$

where $0 \leq \theta \leq 1$ is an unknown parameter. The following 10 independent observations were taken from such a distribution: $\{3,0,2,1,3,2,1,0,2,1\}$.

7 pts (a) Find the method of moments estimator of θ. What is the method of moments estimate?

Solution: We have that

$$\mathbb{E}X = 0\times\frac{2}{3}\theta + 1\times\frac{1}{3}\theta + 2\times\frac{2}{3}(1-\theta) + 3\times\frac{1}{3}(1-\theta) = \frac{\theta}{3}+\frac{4}{3}-\frac{4\theta}{3}+1-\theta = \frac{7-6\theta}{3}.$$

To get the MME we solve $(7-6\theta)/3=\bar{X}$, which gives us the estimator $\hat{\theta}=(7-3\bar{X})/6$. From the data $\bar{X}=15/10$, which plugging into the estimator leads to the estimate $5/12$.

7 pts (b) What is the maximum likelihood estimate of θ?

Solution: The likelihood of one observation is

$$\theta \mapsto \frac{2}{3}\theta 1_{\{X=0\}} + \frac{1}{3}\theta 1_{\{X=1\}} + \frac{2}{3}(1-\theta)1_{\{X=2\}} + \frac{1}{3}(1-\theta)1_{\{X=3\}},$$

which, combined with the fact that we observed $\{3,0,2,\ldots,1\}$ leads to the likelihood,

$$L(\theta) = \frac{1}{3}(1-\theta)\times\frac{2}{3}\theta\times\frac{2}{3}(1-\theta)\times\cdots\times\frac{1}{3}\theta = \left(\frac{2}{3}\theta\right)^2\times\left(\frac{1}{3}\theta\right)^3\times\left(\frac{2}{3}(1-\theta)\right)^3\times\left(\frac{1}{3}(1-\theta)\right)^2$$
$$= \left(\frac{2}{9}\theta(1-\theta)\right)^5.$$

From here it is already clear that this function is maximised by $\tilde{\theta}=1/2$, but this also follows from noting that the log-likelihood is

$$\ell(\theta) = 5\log\frac{2}{9} + 5\log\theta + 5\log(1-\theta);$$

taking derivatives and setting to zero we get

$$\frac{d\ell(\theta)}{d\theta} = \frac{5}{\theta} - \frac{5}{1-\theta} = 0 \Leftrightarrow 1-\theta=\theta \Leftrightarrow \theta = 1/2,$$

which is our maximum likelihood estimate.

7 pts (c) If the prior distribution on θ is $U[0,1]$, what is the posterior density?

Solution: From the previous question we know that the likelihood is

$$L(\theta) = \left(\frac{2}{9}\theta(1-\theta)\right)^5.$$

multiplying this with the prior $\pi(\theta)=1_{\{0\leq\theta<1\}}$ we get that the posterior is proportional to

$$\pi(\theta \mid 3,0,2,\ldots,1) \propto \theta^5(1-\theta)^5 1_{\{0\leq\theta<1\}}.$$

We recognise this as the density of a Beta(6,6) distribution.

4 pts (d) Sketch the posterior. What is the mode of the posterior?

Solution: The mode of the posterior (the MAP estimator) is just 1/2 and coincides with the maximum likelihood estimator since we used a uniform prior. A sketch of the posterior density is depicted in Figure 1.

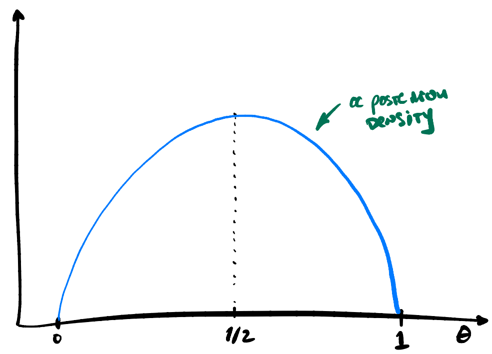

Figure 1: Sketch of the posterior density.

Prob.III: Suppose that you are managing an email mailing list for a website. Let $X_1, \dots, X_n$ be a random sample from the Bernoulli distribution with unknown parameter $p \in [0,1]$. Each observation X_i corresponds to a different person that got the mailing list email. Suppose that $X_i = 1$ if the person opened the email, and $X_i = 0$, otherwise. You are interested in the value of p, the probability that someone who gets the mailing list actually opens it. (The other parameter in the model, n, is not specified but you always know what it is.)

7 pts (a) Describe a near-pivot T for p. Justify how you arrived at T, and don't forget to mention the (approximate) distribution of the pivot.

Solution: Let $\hat{p} = \bar{X}$ be the maximum likelihood estimator for p. The central limit tells us that if n is large enough, then

$$\frac{\bar{X} - \mathbb{E}\bar{X}}{\sqrt{\mathbb{V}\bar{X}}} = \frac{\bar{X} - p}{\sqrt{p(1-p)/n}} = \frac{n\bar{X} - np}{\sqrt{np(1-p)}} \approx N(0,1).$$

The above is already a near-pivot but the pivot can be simplified by noting that the (weak/strong) law of large numbers implies that $\hat{p}$ converges in probability to p so that we are also allowed to claim that

$$T = \frac{n\bar{X} - np}{\sqrt{n\hat{p}(1-\hat{p})}} \approx N(0,1).$$

12 pts (b) Use the near-pivot from the previous question to derive a confidence interval of level 0.95 for p. Don't just present the final result: show how you go from (near-)pivot to confidence interval.

Solution: Since T is a near-pivot and is approximately standard normal distributed, we know that since $z_{0.025} = -z_{0.975}$

$$0.95 \approx \mathbb{P}\left(-z_{0.975} \leq \frac{n\bar{X} - np}{\sqrt{n\hat{p}(1-\hat{p})}} \leq z_{0.975}\right) = \cdots =$$
$$= \mathbb{P}\left(\bar{X} - \frac{\sqrt{n\hat{p}(1-\hat{p})}}{n} z_{0.975} \leq p \leq \bar{X} + \frac{\sqrt{n\hat{p}(1-\hat{p})}}{n} z_{0.975}\right),$$

which leads to a confidence interval for p of level 0.95:

$$\left[\bar{X} - \frac{\sqrt{n\hat{p}(1-\hat{p})}}{n} z_{0.975},\ \bar{X} + \frac{\sqrt{n\hat{p}(1-\hat{p})}}{n} z_{0.975}\right].$$

7 pts (c) Suppose now that the last issue of the mailing list was sent to $n = 27$ subscribers and that, out of these, only 2 opened the email. Based on the confidence interval for p that you found, what would be your most optimistic guess as to how large p is? (There are tables at the end of the exam questionnaire in Appendix A that you may need to consult to answer this question.)

Solution: Using this information we see that $n = 27$, $\hat{p} = 2/27 \approx 0.074$, and from Appendix A we see that $z_{0.975} \approx 1.65$ (read this in the row labeled 1.6 and column

labeled 0.05) . The upper bound of the confidence interval gives us our most optimistic guess on how large p can be:

$$\bar{X} + \frac{\sqrt{n\hat{p}(1-\hat{p})}}{n} z_{0.975} = \frac{2}{27} + \frac{\sqrt{27 \times (2/27) \times (1 - 2/27)}}{27} 1.65 \approx 0.157,$$

which is not very high.

11 pts (d) Suppose that you are not happy with the large amount of uncertainty in the confidence interval that you got – you find the confidence interval too wide. In class, we saw that increasing the sample size n typically reduces uncertainty. How many emails would you have to send to be fairly confident that the estimate $\bar{X}$ of p is not off by more that 0.005? (You can still use $\bar{X}$ as your best guess for p.)

Solution: So that $|\hat{p} - p| = |\bar{X} - p|$ is no more than 0.005 we need to ensure that

$$\frac{\sqrt{n\hat{p}(1-\hat{p})}}{n} z_{0.975} = \sqrt{\frac{\bar{X} \times (1 - \bar{X})}{n}} 1.65 \leq 0.005.$$

Solving for n we conclude that (using $\bar{X} = 2/27 \approx 0.074$)

$$n \geq \frac{1.65^2 \times \bar{X}}{0.005^2} \approx 8066.67,$$

so we conclude that we would need at least (approximately) 8067 observations.

A Quantiles and cumulative probability tables

Below you can find a table of probabilities for the t-distribution and for the standard normal distribution, respectively. Please read the caption of the tables carefully.

$\downarrow \nu/\alpha \rightarrow$	0.3	0.2	0.15	0.1	0.05	0.025	0.02	0.01	0.005	0.0025	0.001
1	0.727	1.376	1.963	3.078	6.314	12.71	15.90	31.82	63.66	127.3	318.3
2	0.617	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.10	22.33
3	0.584	0.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.215
4	0.569	0.941	1.190	1.533	2.132	**2.776**	2.999	3.747	4.604	5.598	7.173
5	0.559	0.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893
6	0.553	0.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208
7	0.549	0.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785
8	0.546	0.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501
9	0.543	0.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297
10	0.542	0.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144
11	0.540	0.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025
12	0.539	0.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930
13	0.538	0.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852
14	0.537	0.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787
15	0.536	0.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733
16	0.535	0.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686
17	0.534	0.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646
18	0.534	0.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.610
19	0.533	0.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579
20	0.533	0.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552
21	0.532	0.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527
22	0.532	0.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505
23	0.532	0.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485
24	0.531	0.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467
25	0.531	0.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450
26	0.531	0.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435
27	0.531	0.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421
28	0.530	0.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408
29	0.530	0.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396
∞	0.524	0.842	1.036	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.090

Figure 2: Right quantiles of the student t-distribution with ν degrees of freedom. Example: if X is a student t distributed random variable, with $\nu = 4$ degrees of freedom then $P(X \geq 2.776) = 0.025$. **The entries of the table are therefore $t_{\nu,1-\alpha}$.**

z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
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	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	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	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	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	0.9875	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	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998
3.6	0.9998	0.9998	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.7	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.8	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999	0.9999
3.9	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

Figure 3: Cumulative probabilities of the standard normal distribution. The rows correspond to the number rounded down to the closest decimal, and columns represent the second decimal. Example: if Z is a standard normal random variable then $P(Z \leq 1.35) = 0.9115$. You look this number up in the row corresponding to 1.3, and in the 6th column (corresponding to 0.05.) Note that for $z < 0$ we have $P(Z \leq z) = 1 - P(Z \leq -z)$. So for example $P(Z \leq -1.35) = 1 - 0.9115 = 0.0885$.