18.650. Fall 2023
Prof. Rigollet

Test 2: review sheet

Here is a list of important results and concepts that can be used in the exam.

1 Parametric inference

1.1 General

- ☐ Parameters of interest vs nuisance parameters
- ☐ Identifiability
- ☐ Estimators

1.2 Method of Moments

- ☐ Moments of a random variable
- ☐ Using plug-in method to generate estimators
- ☐ Consistency and Asymptotic normality, through the delta method

1.3 Maximum Likelihood Estimator

- ☐ Likelihood and log-likelihood
- ☐ The estimator itself – maximizing (log-)likelihood
- ☐ Consistency and Asymptotic normality, under certain regularity conditions
- ☐ Fisher information
- ☐ EM Algorithm - what it does at a high-level (no need to know specifics)

2 Bootstrapping

- ☐ Understanding the emiprical CDF of a sample
- ☐ Understanding how to make bootstrap samples
- ☐ Properties of bootstrap samples (expected value, variance)
- ☐ Purpose of bootstrap (estimate variability, use in confidence intervals, use in tests)
- ☐ Jackknife (difference with Bootstrap)

3 Hypothesis Testing

- ☐ Terminology: test, rejection region, critical value, one-/two-sidedness, level, power
- ☐ Null and Alternate hypotheses
- ☐ Type 1 and type 2 errors
- ☐ p-value and evidence scale
- ☐ Wald test – based off of asymptotically normal estimators

Exercise

Let $X_1, \dots, X_n$ be i.i.d. random variables that follow a Rayleigh distribution with parameter $\theta > 0$. In other words, each X_i has density

$$f_\theta(x) = \frac{x}{\theta^2} \exp\left(-\frac{x^2}{2\theta^2}\right) \mathbf{1}(x \geq 0).$$

It turns out that $\mathbb{E}[X_1^2] = 2\theta^2$.

1. Compute the likelihood $L_n(\theta)$ and the log-likelihood $\ell_n(\theta)$.

   Solution: The likelihood function is

$$\begin{aligned} L_n(X_1, \dots, X_n, \theta) &= \prod_{i=1}^n L(X_i; \theta) \\ &= \frac{1}{\theta^{2n}} (\prod_{i=1}^n X_i) e^{-\frac{\sum_{i=1}^n X_i^2}{2\theta^2}}. \end{aligned}$$

   Thus, our log-likelihood function is

$$\ell_n(X_1, \dots, X_n, \theta) = -2n \log\theta + \sum_{i=1}^n \log X_i - \frac{1}{2\theta^2} \sum_{i=1}^n X_i^2.$$

2. Compute the maximum likelihood estimator $\hat{\theta}^{\mathsf{MLE}}$ for θ.

Solution: Taking the derivative of our log-likehood function, we find that

$$\begin{aligned}
\ell'(\hat{\theta}^{MLE}) &= 0 \\
-\frac{2n}{\hat{\theta}} + \frac{1}{\hat{\theta}^3}\sum_{i=1}^{n} X_i^2 &= 0 \\
\implies \hat{\theta}^2 &= \frac{1}{2n}\sum_{i=1}^{n} X_i^2 \\
\implies \hat{\theta}^{MLE} &= \sqrt{\frac{1}{2n}\sum_{i=1}^{n} X_i^2}.
\end{aligned}$$

3. Verify that $\hat{\theta}^{\mathsf{MLE}}$ is asymptotically normal and determine the asymptotic variance.

   Solution: Since the support of f_θ does not depend on θ and its derivative is zero at θ, this distribution satisfies the conditions required for the MLE to be asymptotically normal. To determine this variance, we first calculate the Fisher information of θ. We have that

$$\begin{aligned}
I(\theta) &= -\mathbb{E}[\ell''(\theta)] \\
&= -\frac{2}{\theta^2} + \frac{3\mathbb{E}[X^2]}{\theta^4} \\
&= -\frac{2}{\theta^2} + \frac{3 * 2\theta^2}{\theta^4} \\
&= -\frac{2}{\theta^2} + \frac{6}{\theta^2} = \frac{4}{\theta^2}.
\end{aligned}$$

   Thus, the asymptotic variance is the inverse of the Fisher information, $\frac{\theta^2}{4}$.

4. Give an asymptotic 95% confidence interval for θ.

Solution: From the previous part, we have that the asymptotic variance of our MLE is $\theta^2/4$. Thus, we have from the asymptotic normality of our MLE that

$$\sqrt{n}\frac{(\hat{\theta}-\theta)}{\sqrt{\theta^2/4}} = \sqrt{n}\frac{(\hat{\theta}-\theta)}{\theta/2} \rightsquigarrow \mathcal{N}(0,1).$$

Using the plug-in method, we replace θ with $\hat{\theta}$ to get that

$$\sqrt{n}\frac{(\hat{\theta}-\theta)}{\hat{\theta}/2} \rightsquigarrow \mathcal{N}(0,1).$$

Thus, our confidence interval is simply

$$(\hat{\theta} - \frac{z_{2.5\%}\hat{\theta}}{2\sqrt{n}}, \hat{\theta} + \frac{z_{2.5\%}\hat{\theta}}{2\sqrt{n}}).$$

We want to test

$$H_0 : \theta = 2 \quad \text{vs.} \quad H_1 : \theta \neq 2.$$

To that end, we collect $n = 400$ observations and compute that $\hat{\theta} = 2.26$.

5. What is the p-value of the Wald test? How strong is the evidence against H_0?

Solution: Under our null hypothesis, $\theta = 2$, so our asymptotic variance is $\frac{\theta^2}{4} = 1$, so we should have $\sqrt{n}(\hat{\theta} - 2) \rightsquigarrow \mathcal{N}(0,1)$. In other words, we should have $(\hat{\theta} - 2) \approx \mathcal{N}(0, \frac{1}{n}) = \mathcal{N}(0, \frac{1}{400})$.

With an observation of $\hat{\theta} = 2.26$, we have a p-value α that satisfies

$$\begin{aligned} |W| &\geq \Phi^{-1}(1-\alpha/2) \\ \left|\frac{2.26-2}{\sqrt{1/400}}\right| &\geq \Phi^{-1}(1-\alpha/2) \\ 5.2 &\geq \Phi^{-1}(1-\alpha/2) \\ \implies \alpha &= 2(1-\Phi(5.2)) \approx 2 \times 10^{-7}. \end{aligned}$$

Thus, we have very strong evidence against H_0.

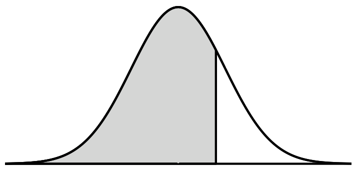

Table 1: The table lists $P(Z \leq z)$ where $Z \sim N(0,1)$ for positive values of z.

	Second decimal place of Z									
Z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	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	0.7704	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

*For $Z \geq 3.50$, the probability is greater than or equal to 0.9998.