

18.05 Practice Exam 2b

No books or calculators. You may have one side of an 8×11 sheet of paper with any information you like on it.
7 problems, 7 pages

Simplifying expressions: You don't need to simplify complicated expressions. For example, you can leave $\frac{1}{4} \cdot \frac{2}{3} + \frac{1}{3} \cdot \frac{2}{5}$ exactly as is. Likewise for expressions like $\frac{20!}{18!2!}$.

The z , t and χ^2 tables are at the end of the exam if you need them.

Problem 1. Concept questions

(a) A certain august journal publishes psychological research. They will only publish results that are statistically significant when tested at a significance level of 0.05.

Could all of their published results be true? **Yes** **No**

Could all of their published results be false? **Yes** **No**

(b) True or false: Setting the prior probability of a hypothesis to 0 means that no amount of data will make the posterior probability of that hypothesis the maximum over all hypotheses.

True **False**

(c) A researcher collected data that fit the criteria for a two-sided Z -test. He set the significance level at 0.05. He ran 80 trials and got a z -value of 1.7. This gave a p -value of 0.0892, so he could not reject the null hypothesis. Convinced that his alternative hypothesis was correct he ran 80 more trials. The combined data from the 160 trials now had a z -value of 2.1. He wrote a paper carefully describing his experiments and submitted it to the journal in part (a).

Will the journal publish his results? **Yes** **No**

(d) Let θ be the probability of heads for a bent coin. Suppose your prior $f(\theta)$ is Beta(6, 8). Also suppose you flip the coin 7 times, getting 2 heads and 5 tails. What is the posterior pdf $f(\theta|x)$?

Problem 2. The Pareto distribution with parameter α has range $[1, \infty)$ and pdf

$$f(x) = \frac{\alpha}{x^{\alpha}}$$

Suppose the data

5, 2, 3

was drawn independently from such a distribution. Find the maximum likelihood estimate (MLE) of α .

Problem 3.

Your friend grabs a die at random from a drawer containing two 4-sided dice, one 8-sided die, and one 12-sided die. They roll the die once and report that the result is 5.

- (a) Make a discrete Bayes table showing the prior, likelihood, and posterior for the type of die rolled given the data.
- (b) What is the prior predictive probability of rolling a 5?
- (c) What are your posterior **odds** that the die has 12 sides?
- (d) Given the data of the first roll, what is your probability that the next roll will be a 7?

Problem 4. Everyone knows that giraffes are tall, but how much do they weigh? Let's suppose that the weight of male giraffes is normally distributed with mean 1200 kg and standard deviation 200 kg.

I volunteered at the zoo and was given the task of weighing their male giraffe Beau. Now weighing a giraffe is not easy and the process produces random errors following a $N(0, 100^2)$ distribution. To compensate for the inaccuracy of the scale I weighed Beau three times and got the following measurements:

1250 kg, 1300 kg, 1350 kg

What is the posterior expected value of Beau's weight?

Problem 5. Data is drawn from a binomial(5, θ) distribution, where θ is unknown. Here is the table of probabilities $p(x | \theta)$ for 3 values of θ :

x	0	1	2	3	4	5
$\theta = 0.5$	0.031	0.156	0.313	0.313	0.156	0.031
$\theta = 0.6$	0.010	0.077	0.230	0.346	0.259	0.078
$\theta = 0.8$	0.000	0.006	0.051	0.205	0.410	0.328

You want to run a significance test on the value of θ . You have the following:

Null hypothesis: $\theta = 0.5$.

Alternate hypotheses: $\theta > 0.5$.

Significance level: $\alpha = 0.1$.

- (a) Find the rejection region.
- (b) Compute the power of the test for each of the two hypotheses $\theta = 0.6$ and $\theta = 0.8$.
- (c) Suppose you run an experiment and the data gives $x = 4$. Compute the p -value of this data.

Problem 6. You have data drawn from a normal distribution with a known variance of 16. You set up the following NHST:

- H_0 : data follows a $N(2, 4^2)$
- H_A : data follows a $N(\mu, 4^2)$ where $\mu \neq 2$.
- Test statistic: standardized sample mean z .
- Significance level set to $\alpha = 0.05$.

You then collected $n = 16$ data points with sample mean 1.5.

- (a) Find the rejection region. Draw a graph indicating the null distribution and the rejection region.
- (b) Find the z -value and add it to your picture in part (a).
- (c) Find the p -value for this data and decide whether or not to reject H_0 in favor of H_A .

Problem 7. Someone claims to have found a long lost work by Jane Austen. She asks you to decide whether or not the book was actually written by Austen.

You buy a copy of *Sense and Sensibility* and count the frequencies of certain common words on some randomly selected pages. You do the same thing for the ‘long lost work’. You get the following table of counts.

Word	a	an	this	that
<i>Sense and Sensibility</i>	150	30	30	90
Long lost work	90	20	10	80

Using this data, set up and evaluate a significance test of the claim that the long lost book is by Jane Austen. Use a significance level of 0.1.

Standard normal table of left tail probabilities.

z	$\Phi(z)$	z	$\Phi(z)$	z	$\Phi(z)$	z	$\Phi(z)$
-4.00	0.0000	-2.00	0.0228	0.00	0.5000	2.00	0.9772
-3.95	0.0000	-1.95	0.0256	0.05	0.5199	2.05	0.9798
-3.90	0.0000	-1.90	0.0287	0.10	0.5398	2.10	0.9821
-3.85	0.0001	-1.85	0.0322	0.15	0.5596	2.15	0.9842
-3.80	0.0001	-1.80	0.0359	0.20	0.5793	2.20	0.9861
-3.75	0.0001	-1.75	0.0401	0.25	0.5987	2.25	0.9878
-3.70	0.0001	-1.70	0.0446	0.30	0.6179	2.30	0.9893
-3.65	0.0001	-1.65	0.0495	0.35	0.6368	2.35	0.9906
-3.60	0.0002	-1.60	0.0548	0.40	0.6554	2.40	0.9918
-3.55	0.0002	-1.55	0.0606	0.45	0.6736	2.45	0.9929
-3.50	0.0002	-1.50	0.0668	0.50	0.6915	2.50	0.9938
-3.45	0.0003	-1.45	0.0735	0.55	0.7088	2.55	0.9946
-3.40	0.0003	-1.40	0.0808	0.60	0.7257	2.60	0.9953
-3.35	0.0004	-1.35	0.0885	0.65	0.7422	2.65	0.9960
-3.30	0.0005	-1.30	0.0968	0.70	0.7580	2.70	0.9965
-3.25	0.0006	-1.25	0.1056	0.75	0.7734	2.75	0.9970
-3.20	0.0007	-1.20	0.1151	0.80	0.7881	2.80	0.9974
-3.15	0.0008	-1.15	0.1251	0.85	0.8023	2.85	0.9978
-3.10	0.0010	-1.10	0.1357	0.90	0.8159	2.90	0.9981
-3.05	0.0011	-1.05	0.1469	0.95	0.8289	2.95	0.9984
-3.00	0.0013	-1.00	0.1587	1.00	0.8413	3.00	0.9987
-2.95	0.0016	-0.95	0.1711	1.05	0.8531	3.05	0.9989
-2.90	0.0019	-0.90	0.1841	1.10	0.8643	3.10	0.9990
-2.85	0.0022	-0.85	0.1977	1.15	0.8749	3.15	0.9992
-2.80	0.0026	-0.80	0.2119	1.20	0.8849	3.20	0.9993
-2.75	0.0030	-0.75	0.2266	1.25	0.8944	3.25	0.9994
-2.70	0.0035	-0.70	0.2420	1.30	0.9032	3.30	0.9995
-2.65	0.0040	-0.65	0.2578	1.35	0.9115	3.35	0.9996
-2.60	0.0047	-0.60	0.2743	1.40	0.9192	3.40	0.9997
-2.55	0.0054	-0.55	0.2912	1.45	0.9265	3.45	0.9997
-2.50	0.0062	-0.50	0.3085	1.50	0.9332	3.50	0.9998
-2.45	0.0071	-0.45	0.3264	1.55	0.9394	3.55	0.9998
-2.40	0.0082	-0.40	0.3446	1.60	0.9452	3.60	0.9998
-2.35	0.0094	-0.35	0.3632	1.65	0.9505	3.65	0.9999
-2.30	0.0107	-0.30	0.3821	1.70	0.9554	3.70	0.9999
-2.25	0.0122	-0.25	0.4013	1.75	0.9599	3.75	0.9999
-2.20	0.0139	-0.20	0.4207	1.80	0.9641	3.80	0.9999
-2.15	0.0158	-0.15	0.4404	1.85	0.9678	3.85	0.9999
-2.10	0.0179	-0.10	0.4602	1.90	0.9713	3.90	1.0000
-2.05	0.0202	-0.05	0.4801	1.95	0.9744	3.95	1.0000

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

(Use interpolation to estimate z values to a 3rd decimal place.)

Table of Student t critical values (right-tail)

The table shows $t_{df,p}$ = the $1 - p$ quantile of $t(df)$.

We only give values for $p \leq 0.5$. Use symmetry to find the values for $p > 0.5$, e.g.

$$t_{5,0.975} = -t_{5,0.025}$$

In R notation $t_{df,p} = \text{qt}(1-p, \text{df})$.

df \ p	0.005	0.010	0.015	0.020	0.025	0.030	0.040	0.050	0.100	0.200	0.300	0.400	0.500
1	63.66	31.82	21.20	15.89	12.71	10.58	7.92	6.31	3.08	1.38	0.73	0.32	0.00
2	9.92	6.96	5.64	4.85	4.30	3.90	3.32	2.92	1.89	1.06	0.62	0.29	0.00
3	5.84	4.54	3.90	3.48	3.18	2.95	2.61	2.35	1.64	0.98	0.58	0.28	0.00
4	4.60	3.75	3.30	3.00	2.78	2.60	2.33	2.13	1.53	0.94	0.57	0.27	0.00
5	4.03	3.36	3.00	2.76	2.57	2.42	2.19	2.02	1.48	0.92	0.56	0.27	0.00
6	3.71	3.14	2.83	2.61	2.45	2.31	2.10	1.94	1.44	0.91	0.55	0.26	0.00
7	3.50	3.00	2.71	2.52	2.36	2.24	2.05	1.89	1.41	0.90	0.55	0.26	0.00
8	3.36	2.90	2.63	2.45	2.31	2.19	2.00	1.86	1.40	0.89	0.55	0.26	0.00
9	3.25	2.82	2.57	2.40	2.26	2.15	1.97	1.83	1.38	0.88	0.54	0.26	0.00
10	3.17	2.76	2.53	2.36	2.23	2.12	1.95	1.81	1.37	0.88	0.54	0.26	0.00
16	2.92	2.58	2.38	2.24	2.12	2.02	1.87	1.75	1.34	0.86	0.54	0.26	0.00
17	2.90	2.57	2.37	2.22	2.11	2.02	1.86	1.74	1.33	0.86	0.53	0.26	0.00
18	2.88	2.55	2.36	2.21	2.10	2.01	1.86	1.73	1.33	0.86	0.53	0.26	0.00
19	2.86	2.54	2.35	2.20	2.09	2.00	1.85	1.73	1.33	0.86	0.53	0.26	0.00
20	2.85	2.53	2.34	2.20	2.09	1.99	1.84	1.72	1.33	0.86	0.53	0.26	0.00
21	2.83	2.52	2.33	2.19	2.08	1.99	1.84	1.72	1.32	0.86	0.53	0.26	0.00
22	2.82	2.51	2.32	2.18	2.07	1.98	1.84	1.72	1.32	0.86	0.53	0.26	0.00
23	2.81	2.50	2.31	2.18	2.07	1.98	1.83	1.71	1.32	0.86	0.53	0.26	0.00
24	2.80	2.49	2.31	2.17	2.06	1.97	1.83	1.71	1.32	0.86	0.53	0.26	0.00
25	2.79	2.49	2.30	2.17	2.06	1.97	1.82	1.71	1.32	0.86	0.53	0.26	0.00
30	2.75	2.46	2.28	2.15	2.04	1.95	1.81	1.70	1.31	0.85	0.53	0.26	0.00
31	2.74	2.45	2.27	2.14	2.04	1.95	1.81	1.70	1.31	0.85	0.53	0.26	0.00
32	2.74	2.45	2.27	2.14	2.04	1.95	1.81	1.69	1.31	0.85	0.53	0.26	0.00
33	2.73	2.44	2.27	2.14	2.03	1.95	1.81	1.69	1.31	0.85	0.53	0.26	0.00
34	2.73	2.44	2.27	2.14	2.03	1.95	1.80	1.69	1.31	0.85	0.53	0.26	0.00
35	2.72	2.44	2.26	2.13	2.03	1.94	1.80	1.69	1.31	0.85	0.53	0.26	0.00
40	2.70	2.42	2.25	2.12	2.02	1.94	1.80	1.68	1.30	0.85	0.53	0.26	0.00
41	2.70	2.42	2.25	2.12	2.02	1.93	1.80	1.68	1.30	0.85	0.53	0.25	0.00
42	2.70	2.42	2.25	2.12	2.02	1.93	1.79	1.68	1.30	0.85	0.53	0.25	0.00
43	2.70	2.42	2.24	2.12	2.02	1.93	1.79	1.68	1.30	0.85	0.53	0.25	0.00
44	2.69	2.41	2.24	2.12	2.02	1.93	1.79	1.68	1.30	0.85	0.53	0.25	0.00
45	2.69	2.41	2.24	2.12	2.01	1.93	1.79	1.68	1.30	0.85	0.53	0.25	0.00
46	2.69	2.41	2.24	2.11	2.01	1.93	1.79	1.68	1.30	0.85	0.53	0.25	0.00
47	2.68	2.41	2.24	2.11	2.01	1.93	1.79	1.68	1.30	0.85	0.53	0.25	0.00
48	2.68	2.41	2.24	2.11	2.01	1.93	1.79	1.68	1.30	0.85	0.53	0.25	0.00
49	2.68	2.40	2.24	2.11	2.01	1.93	1.79	1.68	1.30	0.85	0.53	0.25	0.00

Table of χ^2 critical values (right-tail)

The table shows $c_{df,p}$ = the $1 - p$ quantile of $\chi^2(df)$.

In R notation $c_{df,p} = \text{qchisq}(1-p, df)$.

df\p	0.010	0.025	0.050	0.100	0.200	0.300	0.500	0.700	0.800	0.900	0.950	0.975	0.990
1	6.63	5.02	3.84	2.71	1.64	1.07	0.45	0.15	0.06	0.02	0.00	0.00	0.00
2	9.21	7.38	5.99	4.61	3.22	2.41	1.39	0.71	0.45	0.21	0.10	0.05	0.02
3	11.34	9.35	7.81	6.25	4.64	3.66	2.37	1.42	1.01	0.58	0.35	0.22	0.11
4	13.28	11.14	9.49	7.78	5.99	4.88	3.36	2.19	1.65	1.06	0.71	0.48	0.30
5	15.09	12.83	11.07	9.24	7.29	6.06	4.35	3.00	2.34	1.61	1.15	0.83	0.55
6	16.81	14.45	12.59	10.64	8.56	7.23	5.35	3.83	3.07	2.20	1.64	1.24	0.87
7	18.48	16.01	14.07	12.02	9.80	8.38	6.35	4.67	3.82	2.83	2.17	1.69	1.24
8	20.09	17.53	15.51	13.36	11.03	9.52	7.34	5.53	4.59	3.49	2.73	2.18	1.65
9	21.67	19.02	16.92	14.68	12.24	10.66	8.34	6.39	5.38	4.17	3.33	2.70	2.09
10	23.21	20.48	18.31	15.99	13.44	11.78	9.34	7.27	6.18	4.87	3.94	3.25	2.56
16	32.00	28.85	26.30	23.54	20.47	18.42	15.34	12.62	11.15	9.31	7.96	6.91	5.81
17	33.41	30.19	27.59	24.77	21.61	19.51	16.34	13.53	12.00	10.09	8.67	7.56	6.41
18	34.81	31.53	28.87	25.99	22.76	20.60	17.34	14.44	12.86	10.86	9.39	8.23	7.01
19	36.19	32.85	30.14	27.20	23.90	21.69	18.34	15.35	13.72	11.65	10.12	8.91	7.63
20	37.57	34.17	31.41	28.41	25.04	22.77	19.34	16.27	14.58	12.44	10.85	9.59	8.26
21	38.93	35.48	32.67	29.62	26.17	23.86	20.34	17.18	15.44	13.24	11.59	10.28	8.90
22	40.29	36.78	33.92	30.81	27.30	24.94	21.34	18.10	16.31	14.04	12.34	10.98	9.54
23	41.64	38.08	35.17	32.01	28.43	26.02	22.34	19.02	17.19	14.85	13.09	11.69	10.20
24	42.98	39.36	36.42	33.20	29.55	27.10	23.34	19.94	18.06	15.66	13.85	12.40	10.86
25	44.31	40.65	37.65	34.38	30.68	28.17	24.34	20.87	18.94	16.47	14.61	13.12	11.52
30	50.89	46.98	43.77	40.26	36.25	33.53	29.34	25.51	23.36	20.60	18.49	16.79	14.95
31	52.19	48.23	44.99	41.42	37.36	34.60	30.34	26.44	24.26	21.43	19.28	17.54	15.66
32	53.49	49.48	46.19	42.58	38.47	35.66	31.34	27.37	25.15	22.27	20.07	18.29	16.36
33	54.78	50.73	47.40	43.75	39.57	36.73	32.34	28.31	26.04	23.11	20.87	19.05	17.07
34	56.06	51.97	48.60	44.90	40.68	37.80	33.34	29.24	26.94	23.95	21.66	19.81	17.79
35	57.34	53.20	49.80	46.06	41.78	38.86	34.34	30.18	27.84	24.80	22.47	20.57	18.51
40	63.69	59.34	55.76	51.81	47.27	44.16	39.34	34.87	32.34	29.05	26.51	24.43	22.16
41	64.95	60.56	56.94	52.95	48.36	45.22	40.34	35.81	33.25	29.91	27.33	25.21	22.91
42	66.21	61.78	58.12	54.09	49.46	46.28	41.34	36.75	34.16	30.77	28.14	26.00	23.65
43	67.46	62.99	59.30	55.23	50.55	47.34	42.34	37.70	35.07	31.63	28.96	26.79	24.40
44	68.71	64.20	60.48	56.37	51.64	48.40	43.34	38.64	35.97	32.49	29.79	27.57	25.15
45	69.96	65.41	61.66	57.51	52.73	49.45	44.34	39.58	36.88	33.35	30.61	28.37	25.90
46	71.20	66.62	62.83	58.64	53.82	50.51	45.34	40.53	37.80	34.22	31.44	29.16	26.66
47	72.44	67.82	64.00	59.77	54.91	51.56	46.34	41.47	38.71	35.08	32.27	29.96	27.42
48	73.68	69.02	65.17	60.91	55.99	52.62	47.34	42.42	39.62	35.95	33.10	30.75	28.18
49	74.92	70.22	66.34	62.04	57.08	53.67	48.33	43.37	40.53	36.82	33.93	31.55	28.94

MIT OpenCourseWare

<https://ocw.mit.edu>

18.05 Introduction to Probability and Statistics

Spring 2022

For information about citing these materials or our Terms of Use, visit: <https://ocw.mit.edu/terms>.