Midterm Examination

Mth 136 = Sta 114

Wednesday, 2001 April 11, 2:20 – 3:35 pm

This is a closed-book examination so please do not refer to your notes, the text, or to any other books. You may use a one-sided single sheet of *your own* notes, if you wish, but you may not share materials. Distribution tables for the normal, t, and chi-squared distributions, the PDF handout, and a blank worksheet are attached to the exam. If you don't understand something in one of the tables or questions feel free to ask me, but please do not talk to each other. You may use a calculator but not a laptop computer.

You must **show your work** to get partial credit. Unsupported answers are not acceptable, even if they are correct. Please give all numerical answers as fractions **in lowest terms** (simplify!) or as decimals **correct to four places**. You should spend about 10–15 minutes on each problem. It is to your advantage to write your solutions as clearly as possible, since I cannot give credit for solutions I do not understand. Good luck.

Cheating on exams is a breach of trust with classmates and faculty, and will not be tolerated. After completing the exam please acknowledge the Duke Honor Code:

I have neither given nor received any unauthorized aid on this exam.

Signature: ______________________

Print Name: ______________________

1.	/20
2.	/20
3.	/20
4.	/20
5.	/20
Total:	/100

1. Skyler is following a new stock, XYZ, and hopes for frequent 1% daily gains. Let θ be the probability of a 1% gain, and assume that daily gains are independent so that the events $E_i : \{ \text{ 1\% gain on day } i \}$ are independent, each with probability θ. To test the hypothesis

$$\begin{aligned} H_0 &: \quad \theta = 0.20 \\ H_1 &: \quad \theta = 0.10 \end{aligned}$$

Skyler decides to count the number X of days preceding (and not including) the first day on which XYZ gains 1% (so $X \geq 0$ has a geometric distribution starting at zero).

a. (2pt) As a function of $\theta > 0$ and $k \in \mathbb{N}$, give the indicated probability. Hint: *Either* evaluate the sum of a geometric series *or* think through what must happen on the first k trials.
$\mathsf{P}[X \geq k \mid \theta] =$ ______________

b. (8pt) If Skyler decides to Reject H_0 if $X \geq 10$, determine the probabilities of Type-I and Type-II errors. Remember the instructions about simplifying numeric answers.
$\alpha =$ ______________ $\beta =$ ______________

c. (10pt) If Skyler assigns equal prior probability $\xi(.20) = \xi(.10) = .5$, for which X is the posterior probability of H_0 greater than one-half?
$\Pr[H_0 \mid X] > 0.50$ for: __________ $\leq X \leq$ __________

2. A box contains a large number of chips of three different colors, red brown and blue; it is desired to test the hypothesis that chips of the three colors are present in equal proportions, against the alternative that they are not. A draw of thirty chips produces **8 red**, **15 brown**, and **7 blue** chips.

a. (10pt) What statistical procedure would be most appropriate?

○ t test ○ Normal (z) test ○ χ^2 test ○ Confidence Interval
○ Simple linear regression ○ Multiple linear regression ○ Other

Degrees of Freedom (if any): $\nu =$ __________

b. (10pt) Perform the procedure and report your findings (interpolate or approximate if necessary to find the p-value).
$p \approx$ __________
Choose one: at $\alpha = .10$, ○ Reject ○ Accept

3. A **single observation** X is to be drawn from pdf

$$\begin{aligned} f(x|\theta) &= 2(1-\theta)x+\theta, \quad 0<x<1 \\ &= (1-2x)\theta+2x, \quad 0<x<1 \end{aligned}$$

for some unknown θ in the range $0 \le \theta \le 2$.

a. (10pt) Find the Maximum Likelihood Estimator of θ, as a function of the **single observation** x (Hint: sketch the likelihood for a few values of x):

$\hat{\theta}(x) =$

b. (10pt) Test the hypothesis

$$H_0 : [\theta = 1] \quad vs. \quad H_1 : [\theta < 1]$$

for the observation $X = 0.25$— i.e., find the p-value and decide whether to accept or not. You may pick the size of the test.

$p =$ __________

Choose one: at $\alpha =$ __________, ○ Reject ○ Accept

4. In forestry, the diameter of a tree at breast height (DBH), which is fairly easy to measure, is used to predict the height of the tree, which is more difficult to measure. Silviculturalists working in British Columbia's boreal forest conducted a series of spacing trials to predict the heights of several species of trees. The data in the accompanying table are the breast height diameters (in centimeters) and heights (in meters) of 36 white spruce trees (*Please* don't do any arithmetic with these):

Diam (cm)	18.9	15.5	19.4	20.0	29.8	19.8	20.3	20.0	22.0	23.6	14.8	22.7
Height (m)	20.0	16.8	20.2	20.0	20.2	18.0	17.8	19.2	22.3	18.9	13.3	20.6
Diam (cm)	18.5	21.5	14.8	17.7	21.0	15.9	16.6	15.5	13.7	27.5	20.3	22.9
Height (m)	19.0	19.2	16.1	19.9	20.4	17.6	18.8	16.9	16.3	21.4	19.2	19.8
Diam (cm)	14.1	10.1	5.8	20.7	17.8	11.4	14.4	13.4	17.8	20.7	13.3	22.9
Height (m)	18.5	12.1	8.0	17.4	18.4	17.3	16.6	12.9	17.5	19.4	15.5	19.2

a. (6pt) What statistical procedure would be most appropriate?

○ t test ○ Normal (z) test ○ χ^2 test ○ Confidence Interval
○ Simple linear regression ○ Multiple linear regression ○ Other
Degrees of Freedom (if any): $\nu =$ __________

Here is an S-Plus output for the linear model function `lm()`:

```
Call:  lm(formula = height ~ diam, data = spruce)
Residuals:
   Min       1Q  Median     3Q    Max
-3.939  -0.9763  0.2829  0.995  2.664

 Coefficients:Value  Std.  Error  t value  Pr(>|t|)
(Intercept)  9.1468       1.1213   8.1573    0.0000
       diam  0.4815       0.0597   8.0685    0.0000

Residual standard error:  1.678 on 34 degrees of freedom
Multiple R-Squared:  0.6569
F-statistic:  65.1 on 1 and 34 degrees of freedom, the p-value is 2.089e-009
```

b. (4pt) Assuming the relationship between the variables is best described by a straight line, estimate the regression equation:

$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X =$ __________ + __________ X

c. (10pt) Give a 90% confidence interval for the true slope in this equation:

$\beta_1 =$ _______________ ± _______________

5. Recall the spruce data and S-Plus output for the `lm()` function:

```
Call:  lm(formula = height ~ diam, data = spruce)
Residuals:
   Min       1Q  Median     3Q    Max
-3.939  -0.9763  0.2829  0.995  2.664

 Coefficients:Value  Std.  Error  t value  Pr(>|t|)
(Intercept)  9.1468       1.1213   8.1573    0.0000
       diam  0.4815       0.0597   8.0685    0.0000

Residual standard error:  1.678 on 34 degrees of freedom
Multiple R-Squared:  0.6569
F-statistic:  65.1 on 1 and 34 degrees of freedom, the p-value is 2.089e-009
```

a. (10pt) Do the breast height diameters contribute useful information for predicting tree height? Answer this scientific question by testing a suitable Hypothesis:

H_0 : ________________

H_1 : ________________

At $\alpha = .01$, ○ Accept H_0 ○ Reject H_0

Scientific conclusion:

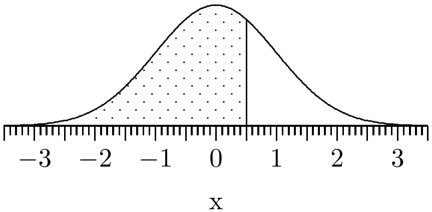
b. (4pt) How tall do you expect a Spruce would be with a DBH of 5cm?

$\mathsf{E}[Y|X=5] \approx$ __________

c. (6pt) How reliable do you think this prediction is? Why? (in words, no calculation is needed)

Extra worksheet, if needed:

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}}\, e^{-z^2/2}\, dz:$$



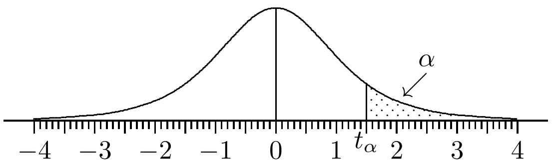
Normal Distribution Table Area $\Phi(x)$ to the left of x.

x	.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

$\Phi(0.6745) = 0.75$ $\Phi(1.6449) = 0.95$ $\Phi(2.3263) = 0.99$ $\Phi(3.0902) = 0.999$

$\Phi(1.2816) = 0.90$ $\Phi(1.9600) = 0.975$ $\Phi(2.5758) = 0.995$ $\Phi(3.2905) = 0.9995$

$$\alpha = \int_{t_\alpha}^{\infty} \frac{c\;dt}{(1+t^2/\nu)^{(\nu+1)/2}}$$



Student's t Distribution Table

ν	$t_{.40}$	$t_{.30}$	$t_{.20}$	$t_{.15}$	$t_{.10}$	$t_{.05}$	$t_{.025}$	$t_{.01}$	$t_{.005}$	$t_{.001}$	$t_{.0005}$	$t_{.0001}$
1	0.325	0.727	1.376	1.9626	3.078	6.314	12.76	31.82	63.66	318.3	636.6	3183.
2	0.289	0.617	1.061	1.3862	1.886	2.920	4.303	6.965	9.925	22.33	31.60	70.70
3	0.277	0.584	0.978	1.2498	1.638	2.353	3.182	4.541	5.841	10.22	12.92	22.20
4	0.271	0.569	0.941	1.1896	1.533	2.132	2.776	3.747	4.604	7.173	8.610	13.03
5	0.267	0.559	0.920	1.1558	1.476	2.015	2.571	3.365	4.032	5.893	6.869	9.678
6	0.265	0.553	0.906	1.1342	1.440	1.943	2.447	3.143	3.707	5.208	5.959	8.025
7	0.263	0.549	0.896	1.1192	1.415	1.895	2.365	2.998	3.499	4.785	5.408	7.063
8	0.262	0.546	0.889	1.1081	1.397	1.860	2.306	2.896	3.355	4.501	5.041	6.442
9	0.261	0.543	0.883	1.0997	1.383	1.833	2.262	2.821	3.250	4.297	4.781	6.010
10	0.260	0.542	0.879	1.0931	1.372	1.812	2.228	2.764	3.169	4.144	4.587	5.694
11	0.260	0.540	0.876	1.0877	1.363	1.796	2.201	2.718	3.106	4.025	4.437	5.453
12	0.259	0.539	0.873	1.0832	1.356	1.782	2.179	2.681	3.055	3.930	4.318	5.263
13	0.259	0.538	0.870	1.0795	1.350	1.771	2.160	2.650	3.012	3.852	4.221	5.111
14	0.258	0.537	0.868	1.0763	1.345	1.761	2.145	2.624	2.977	3.787	4.140	4.985
15	0.258	0.536	0.866	1.0735	1.341	1.753	2.131	2.602	2.947	3.733	4.073	4.880
16	0.258	0.535	0.865	1.0711	1.337	1.746	2.120	2.583	2.921	3.686	4.015	4.791
17	0.257	0.534	0.863	1.0690	1.333	1.740	2.110	2.567	2.898	3.646	3.965	4.714
18	0.257	0.534	0.862	1.0672	1.330	1.734	2.101	2.552	2.878	3.610	3.922	4.648
19	0.257	0.533	0.861	1.0655	1.328	1.729	2.093	2.539	2.861	3.579	3.883	4.590
20	0.257	0.533	0.860	1.0640	1.325	1.725	2.086	2.528	2.845	3.552	3.85	4.539
21	0.257	0.532	0.859	1.0627	1.323	1.721	2.080	2.518	2.831	3.527	3.819	4.493
22	0.256	0.532	0.858	1.0614	1.321	1.717	2.074	2.508	2.819	3.505	3.792	4.452
23	0.256	0.532	0.858	1.0603	1.319	1.714	2.069	2.500	2.807	3.485	3.768	4.415
24	0.256	0.531	0.857	1.0593	1.318	1.711	2.064	2.492	2.797	3.467	3.745	4.382
25	0.256	0.531	0.856	1.0584	1.316	1.708	2.060	2.485	2.787	3.450	3.725	4.352
26	0.256	0.531	0.856	1.0575	1.315	1.706	2.056	2.479	2.779	3.435	3.707	4.324
27	0.256	0.531	0.855	1.0567	1.314	1.703	2.052	2.473	2.771	3.421	3.690	4.299
28	0.256	0.530	0.855	1.0560	1.313	1.701	2.048	2.467	2.763	3.408	3.674	4.275
29	0.256	0.530	0.854	1.0553	1.311	1.699	2.045	2.462	2.756	3.396	3.659	4.254
30	0.256	0.530	0.854	1.0547	1.310	1.697	2.042	2.457	2.750	3.385	3.646	4.234
40	0.255	0.529	0.851	1.0500	1.303	1.684	2.021	2.423	2.704	3.307	3.551	4.094
60	0.254	0.527	0.848	1.0455	1.296	1.671	2.000	2.390	2.660	3.232	3.460	3.962
120	0.254	0.526	0.845	1.0409	1.289	1.658	1.980	2.358	2.617	3.160	3.373	3.837
∞	0.253	0.524	0.842	1.0364	1.282	1.645	1.960	2.326	2.576	3.090	3.291	3.719

$$\alpha = \int_{\chi^2_\alpha}^{\infty} c\,x^{\nu/2-1}e^{-x/2}\,dx$$

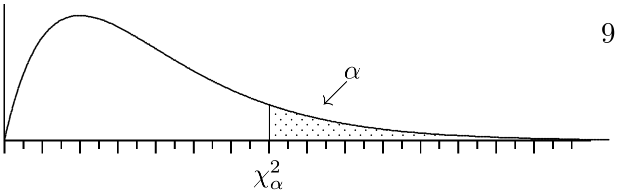


Table of the χ^2 Distribution Area α under the χ^2 Curve to the right of χ^2_α.

ν	$\chi^2_{.50}$	$\chi^2_{.25}$	$\chi^2_{.10}$	$\chi^2_{.05}$	$\chi^2_{.025}$	$\chi^2_{.01}$	$\chi^2_{.005}$	$\chi^2_{.001}$	$\chi^2_{.0005}$	$\chi^2_{.0001}$
1	0.4549	1.3233	2.7055	3.8415	5.0239	6.6349	7.87940	10.8276	12.1157	15.1367
2	1.3863	2.7726	4.6052	5.9915	7.3778	9.2103	10.5966	13.8155	15.2018	18.4207
3	2.3660	4.1083	6.2514	7.8147	9.3484	11.3449	12.8382	16.2662	17.7300	21.1075
4	3.3567	5.3853	7.7794	9.4877	11.1433	13.2767	14.8603	18.4668	19.9974	23.5127
5	4.3515	6.6257	9.2364	11.0705	12.8325	15.0863	16.7496	20.5150	22.1053	25.7448
6	5.3481	7.8408	10.6446	12.5916	14.4494	16.8119	18.5476	22.4577	24.1028	27.8563
7	6.3458	9.0371	12.0170	14.0671	16.0128	18.4753	20.2777	24.3219	26.0178	29.8775
8	7.3441	10.219	13.3616	15.5073	17.5345	20.0902	21.9550	26.1245	27.8680	31.8276
9	8.3428	11.389	14.6837	16.9190	19.0228	21.6660	23.5894	27.8772	29.6658	33.7199
10	9.3418	12.549	15.9872	18.3070	20.4831	23.2092	25.1882	29.5883	31.4198	35.5640
11	10.341	13.701	17.2750	19.6751	21.9200	24.7249	26.7568	31.2641	33.1366	37.3670
12	11.340	14.845	18.5493	21.0260	23.3366	26.2169	28.2995	32.9095	34.8213	39.1344
13	12.340	15.984	19.8119	22.3620	24.7356	27.6882	29.8195	34.5282	36.4778	40.8707
14	13.339	17.117	21.0641	23.6848	26.1189	29.1412	31.3193	36.1233	38.1094	42.5793
15	14.339	18.245	22.3071	24.9958	27.4884	30.5779	32.8013	37.6973	39.7188	44.2632
16	15.338	19.369	23.5418	26.2962	28.8453	31.9999	34.2672	39.2524	41.3081	45.9249
17	16.338	20.489	24.7690	27.5871	30.1910	33.4087	35.7185	40.7902	42.8792	47.5664
18	17.338	21.605	25.9894	28.8693	31.5264	34.8053	37.1565	42.3124	44.4338	49.1894
19	18.338	22.718	27.2036	30.1435	32.8523	36.1909	38.5823	43.8202	45.9731	50.7955
20	19.337	23.828	28.4120	31.4104	34.1696	37.5662	39.9968	45.3147	47.4985	52.3860
21	20.337	24.935	29.6151	32.6706	35.4789	38.9322	41.4011	46.7970	49.0108	53.9620
22	21.337	26.039	30.8133	33.9244	36.7807	40.2894	42.7957	48.2679	50.5111	55.5246
23	22.337	27.141	32.0069	35.1725	38.0756	41.6384	44.1813	49.7282	52.0002	57.0746
24	23.337	28.241	33.1962	36.4150	39.3641	42.9798	45.5585	51.1786	53.4788	58.6130
25	24.337	29.339	34.3816	37.6525	40.6465	44.3141	46.9279	52.6197	54.9475	60.1403
26	25.336	30.435	35.5632	38.8851	41.9232	45.6417	48.2899	54.0520	56.4069	61.6573
27	26.336	31.528	36.7412	40.1133	43.1945	46.9629	49.6449	55.4760	57.8576	63.1645
28	27.336	32.620	37.9159	41.3371	44.4608	48.2782	50.9934	56.8923	59.3000	64.6624
29	28.336	33.711	39.0875	42.5570	45.7223	49.5879	52.3356	58.3012	60.7346	66.1517
30	29.336	34.800	40.2560	43.7730	46.9792	50.8922	53.6720	59.7031	62.1619	67.6326
40	39.336	45.616	51.8051	55.7585	59.3417	63.6907	66.7660	73.4020	76.0946	82.0623
50	49.335	56.334	63.1671	67.5048	71.4202	76.1539	79.4900	86.6608	89.5605	95.9687
60	59.335	66.981	74.3970	79.0819	83.2977	88.3794	91.9517	99.6072	102.695	109.503
70	69.335	77.577	85.5270	90.5312	95.0232	100.425	104.215	112.317	115.578	122.755
80	79.334	88.130	96.5782	101.879	106.629	112.329	116.321	124.839	128.261	135.782
90	89.334	98.650	107.565	113.145	118.136	124.116	128.299	137.208	140.782	148.627
100	99.334	109.14	118.498	124.342	129.561	135.807	140.169	149.449	153.167	161.319