Statistics Qualifying Exam

9:00 am - 1:00 pm, Thursday, February 4, 2016

1. Let X_1 and X_2 have the joint probability mass function (pmf) $p(x_1, x_2)$ described as follows: and $p(x_1, x_2)$

Table 1: Joint pmf of X_1 and X_2

		Support of X_2	
		0	1
Support of X_1	0	$\frac{1}{15}$	$\frac{3}{15}$
	1	$\frac{2}{15}$	$\frac{4}{15}$
	2	$\frac{1}{15}$	$\frac{4}{15}$

is zero elsewhere.

(a) Calculate the conditional mean of X_1 given $X_2 = 1$?.

(b) Calculate the covaraince of X_1, X_2, $Cov(X_1, X_2)$.

(c) Find the conditional probability mass function (pmf) of $p(X_2|X_1 = 1)$

2. Let $X_1 \sim Gamma(2, 2)$, $X_2 \sim Gamma(4, 2)$ are two independent random variables. The probability density function (pdf) of a gamma distribution $Gamma(\alpha, \beta)$ is:

$$f(x) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} \exp\left(-\frac{x}{\beta}\right)$$

(a) Find the mean of $Y = X_1 + 3X_1X_2^2 + 1$. $(E(Y))$

(b) Find the variance of $Z = X_1 + X_2$.

(c) If $Y_1 = X_1 + X_2$ and $Y_2 = X_1$ find the joint pdf of Y_1 and Y_2. Are Y_1 and Y_2 independent?

(d) Find the conditional mean of $Y_1|Y_2 = y_2$.

3. Let $\overline{X}_n$ denote the mean of a random sample of size n from a Poisson distribution with parameter $\mu = 1$.

(a) Show that the mgf of $Y_n = \sqrt{n}(\overline{X}_n - 1)$ is given by $\exp[-t\sqrt{n} + n(e^{t/\sqrt{n}} - 1)]$.

(b) Investigate the limiting distribution of Y_n as $n \to \infty$.

(c) Find the limiting distribution of $\sqrt{n}(\sqrt{\overline{X}_n} - 1)$.

4. Let $\overline{X}$ be the mean of a random sample of size n from $N(\theta, \sigma^2)$ distribution, $-\infty < \theta < \infty$, $\sigma^2 > 0$. Assume that σ^2 is known. Show that $\overline{X}^2 - \frac{\sigma^2}{n}$ is an unbiased estimator of θ^2 and find its efficiency.

5. An investigation is conducted to study gasoline mileage in automobiles when used exclusively for urban driving. Ten properly tuned and serviced automobiles manufactured during the same year are used in the study. Each automobile is driven for 1000 miles, and the average number of miles per gallon (mi/gal) obtained (Y) and the weight of the car in tons (X) are recorded. These data result:

Car number	1	2	3	4	5	6	7	8	9	10
Miles per gallon(y)	17.9	16.5	16.4	16.8	18.8	15.5	17.5	16.4	15.9	18.3
Weight in tons(x)	1.35	1.90	1.70	1.80	1.30	2.05	1.60	1.80	1.85	1.40

Summary statistics for these data are
$\sum x = 16.75$, $\sum x^2 = 28.6375$, $\sum y = 170.0$, $\sum y^2 = 2900.46$, $\sum xy = 282.405$

(a) Fill in the ANOVA table below for the linear regression model:

Source	df	Sum of Square (SS)	Mean Squares (MS)	F-ratio
Regression				
Error				
Total (corrected for mean)				

(b) Give a 95% confidence interval for the slope parameter.

6. A consumer organization studied the effect of age of automobile owner on size of cash offer for a used car by utilizing 12 persons in each of the two age groups (young, middle) who acted as the owner of a used car. A medium price, six-year-old car was selected for the experiment, and the "owners" solicited cash offers for this car from 24 dealers selected at random from the dealers in the region. Randomization was used in assigning the dealers to the "owners." The offers(in hundred dollars) follow.

i \ j	1	2	3	4	5	6	7	8	9	10	11	12		
Young	23	25	21	22	21	22	20	23	19	22	19	21	$\bar{Y}_{1.} = 21.5$	$S_1^2 = 1.73^2$
Middle	28	27	27	29	26	29	27	30	28	27	26	29	$\bar{Y}_{2.} = 27.75$	$S_2^2 = 1.29^2$
													$\bar{Y}_{..} = 24.625$	$S^2 = 3.52^2$

Note: $\bar{Y}_{1.}, \bar{Y}_{2.}, S_1^2$, and S_2^2 represent the sample means and the unbiased sample variances for the age group "young" and "middle", respectively. $\bar{Y}_{..}$ and S^2 represent the grand sample mean and unbiased sample variance.

(a) Use a two-sample t-test to test $H_0 : \mu_1 = \mu_2$ versus $H_0 : \mu_1 \neq \mu_2$, where μ_1 and μ_2 represent the mean size of cash offers for age group "young" and "middle", respectively. Please clearly specify the assumptions made in the procedure. Use $\alpha = 0.05$.

(b) Assume the ANOVA model $Y_{ij} = \mu + \alpha_i + \epsilon_{ij}$ applicable to the above data, where ϵ_{ij} are independent and identical random variables with $\epsilon_{ij} \sim N(0, \sigma^2)$, $i = 1, 2$, and $j = 1, \ldots, 12$. Complete the ANOVA table below and conduct the F test for the equality of factor level means. Clearly state the value of the test statistic, the sampling distribution under the null hypothesis, and your conclusion. Use $\alpha = 0.05$.

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Age Group		-----------	-----------	------	<.0001
Error		-----------	-----------		
Corrected Total		285.6250000			

7. An engineer is designing a battery for use in a device that will be subject to some extreme variations in temperatures. She decides to test three plate materials at three temperature levels. Because there are two factors at three levels, this design is sometimes called a 3^2 factorial design. Four batteries are tested at each combination of plate materials and temperature, and all 36 tests are run in random order. The experiment and the resulting observed battery life data are given in the table below. A longer life is preferred. The overall mean battery life of the sample is 105.53.

Material Type	Temperature(°F)		
	15	70	125
1	130, 155 74, 180	34, 40 80, 75	20, 70 82, 58
2	150, 188 159, 126	136, 122 106, 115	25, 70 58, 45
3	138, 110 168, 160	174, 120 150, 139	96, 104 82, 60

Part of SAS output is included from a two-factor fixed effects ANOVA analysis.

(a) Construct the ANOVA table based on the output from SAS.

  i. Clearly specify the sources of sum of squares, the degrees of freedom, the mean squares, and the values of F statistics.

  ii. State your findings. Use $\alpha = 0.05$ for each F test.

(b) Now assume it is given that $Temperature = 70°F$, Carry out Tukey multiple comparison on the material types effect. Use $\alpha = 0.05$. If the exact critical value needed cannot be found in the statistical tables provided, please choose the most appropriate approximate value and briefly discuss how the approximation affects the decision.

Problem 8 SAS OUTPUT

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	***	59416.22222	7427.02778	***	<.0001
Error	***	***	***		
Corrected Total	***	77646.97222			

Source	DF	Type III SS	Mean Square	F Value	Pr > F
temp	***	39118.72222	***	***	<.0001
type	***	10683.72222	***	***	0.0020
temp*type	***	****	***	***	0.0186

Level of temp	N	life Mean	life Std Dev
15	12	144.833333	31.6940870
70	12	107.583333	42.8834750
125	12	64.166667	25.6721757

Level of type	N	life Mean	life Std Dev
1	12	83.166667	48.5888751
2	12	108.333333	49.4723676
3	12	125.083333	35.7655455

Level of temp	Level of type	N	life Mean	life Std Dev
15	1	4	134.750000	45.3532432
15	2	4	155.750000	25.6173769
15	3	4	144.000000	25.9743463
70	1	4	57.250000	23.5990819
70	2	4	119.750000	12.6589889
70	3	4	145.750000	22.5444006
125	1	4	57.500000	26.8514432
125	2	4	49.500000	19.2613603
125	3	4	85.500000	19.2786583

8. The article "The New Mantra: MVT" (*Forbes*, March 11, 1996, by Koselka, Rita) provides an interesting example on experimental design for a movie theater. Imagine the owner of the movie theater would like to maximize her weekly profit. Her options include: (A) Jack up the ticket price by a buck (B) Take out bigger ads in the local paper (C) Give away the popcorn.

   (a) The owner would like to test these three options in isolation. As a statistician, you need to persuade her to test all three at once. Please list your arguments.

   (b) Assume the experiment is carried out as you suggested (all three at once). The data obtained is as follows.

	A	B	C	
Test	Raise ticket price	Advertise	Give out free popcorn	Profit($ thousand)
1	NO	NO	NO	10
2	NO	NO	YES	15
3	NO	YES	NO	5
4	NO	YES	YES	10
5	YES	NO	NO	12
6	YES	NO	YES	20
7	YES	YES	NO	7
8	YES	YES	YES	15

   Please prepare a mini-report for this data to address the following two questions

   i. What are the estimated effects for the three options (A) Jack up the ticket price by a buck (B) Take out bigger ads in the local paper (C) Give away the popcorn?

   ii. What would you recommend to the owner of the movie theater? Note that since your report is going to be the reviewed by a statistician, simply list the seemingly "best" option (for example, raise ticket price + give out free popcorn) is not going to earn you any credit. Your report needs to include clear description of your statistical model and inference procedures along with sufficient statistical evidence to support your recommendation. In case that you need to calculate the sum of squares (SS) for a factor, the formula is $(contrast)^2/(2^k \cdot n)$, where k is the number of main factors and n is the number of replicates.

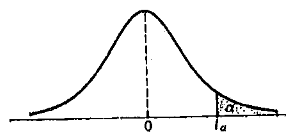


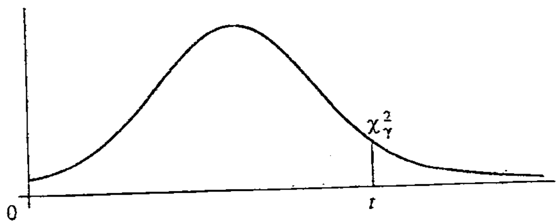
Table A.4 Critical Values of the t-Distribution

	α						
v	**0.40**	**0.30**	**0.20**	**0.15**	**0.10**	**0.05**	**0.025**
1	0.325	0.727	1.376	1.963	3.078	6.314	12.706
2	0.289	0.617	1.061	1.386	1.886	2.920	4.303
3	0.277	0.584	0.978	1.250	1.638	2.353	3.182
4	0.271	0.569	0.941	1.190	1.533	2.132	2.776
5	0.267	0.559	0.920	1.156	1.476	2.015	2.571
6	0.265	0.553	0.906	1.134	1.440	1.943	2.447
7	0.263	0.549	0.896	1.119	1.415	1.895	2.365
8	0.262	0.546	0.889	1.108	1.397	1.860	2.306
9	0.261	0.543	0.883	1.100	1.383	1.833	2.262
10	0.260	0.542	0.879	1.093	1.372	1.812	2.228
11	0.260	0.540	0.876	1.088	1.363	1.796	2.201
12	0.259	0.539	0.873	1.083	1.356	1.782	2.179
13	0.259	0.538	0.870	1.079	1.350	1.771	2.160
14	0.258	0.537	0.868	1.076	1.345	1.761	2.145
15	0.258	0.536	0.866	1.074	1.341	1.753	2.131
16	0.258	0.535	0.865	1.071	1.337	1.746	2.120
17	0.257	0.534	0.863	1.069	1.333	1.740	2.110
18	0.257	0.534	0.862	1.067	1.330	1.734	2.101
19	0.257	0.533	0.861	1.066	1.328	1.729	2.093
20	0.257	0.533	0.860	1.064	1.325	1.725	2.086
21	0.257	0.532	0.859	1.063	1.323	1.721	2.080
22	0.256	0.532	0.858	1.061	1.321	1.717	2.074
23	0.256	0.532	0.858	1.060	1.319	1.714	2.069
24	0.256	0.531	0.857	1.059	1.318	1.711	2.064
25	0.256	0.531	0.856	1.058	1.316	1.708	2.060
26	0.256	0.531	0.856	1.058	1.315	1.706	2.056
27	0.256	0.531	0.855	1.057	1.314	1.703	2.052
28	0.256	0.530	0.855	1.056	1.313	1.701	2.048
29	0.256	0.530	0.854	1.055	1.311	1.699	2.045
30	0.256	0.530	0.854	1.055	1.310	1.697	2.042
40	0.255	0.529	0.851	1.050	1.303	1.684	2.021
60	0.254	0.527	0.848	1.045	1.296	1.671	2.000
120	0.254	0.526	0.845	1.041	1.289	1.658	1.980
∞	0.253	0.524	0.842	1.036	1.282	1.645	1.960

50	15664	10493	20492	38391	91132	21999	59516	81652	27195	48223	46751	22923	32261	85653

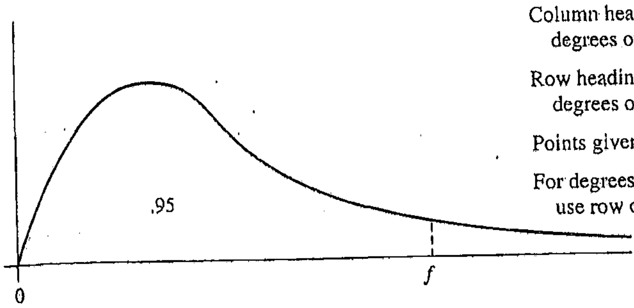
TABLE IV

Cumulative chi-squared distribution



F / γ	$P[\chi^2_\gamma \le t]$												
	0.005	0.010	0.025	0.050	0.100	0.250	0.500	0.750	0.900	0.950	0.975	0.990	0.995
1	0.0000393	0.000157	0.000982	0.00393	0.0158	0.102	0.455	1.32	2.71	3.84	5.02	6.63	7.88
2	0.0100	0.0201	0.0506	0.103	0.211	0.575	1.39	2.77	4.61	5.99	7.38	9.21	10.6
3	0.0717	0.115	0.216	0.352	0.584	1.21	2.37	4.11	6.25	7.81	9.35	11.3	12.8
4	0.207	0.297	0.484	0.711	1.06	1.92	3.36	5.39	7.78	9.49	11.1	13.3	14.9
5	0.412	0.554	0.831	1.15	1.61	2.67	4.35	6.63	9.24	11.1	12.8	15.1	16.7
6	0.676	0.872	1.24	1.64	2.20	3.45	5.35	7.84	10.6	12.6	14.4	16.8	18.5
7	0.989	1.24	1.69	2.17	2.83	4.25	6.35	9.04	12.0	14.1	16.0	18.5	20.3
8	1.34	1.65	2.18	2.73	3.49	5.07	7.34	10.2	13.4	15.5	17.5	20.1	22.0
9	1.73	2.09	2.70	3.33	4.17	5.90	8.34	11.4	14.7	16.9	19.0	21.7	23.6
10	2.16	2.56	3.25	3.94	4.87	6.74	9.34	12.5	16.0	18.3	20.5	23.2	25.2
11	2.60	3.05	3.82	4.57	5.58	7.58	10.3	13.7	17.3	19.7	21.9	24.7	26.8
12	3.07	3.57	4.40	5.23	6.30	8.44	11.3	14.8	18.5	21.0	23.3	26.2	28.3

TABLE IX
F distribution *(continued)*



Column heading = numerator degrees of freedom

Row heading = denominator degrees of freedom

Points given are $f_{.05}$ points

For degrees of freedom > 120, use row or column 120

$P[F_{\gamma_1, \gamma_2} \le f] = .95$

γ_2 \ γ_1	1	2	3	4	5	6	7	8
1	161.448	199.500	215.707	224.583	230.161	233.985	236.768	238.882
2	18.513	19.000	19.164	19.247	19.296	19.329	19.353	19.371
3	10.128	9.552	9.277	9.117	9.013	8.941	8.887	8.845
4	7.709	6.944	6.591	6.388	6.256	6.163	6.094	6.041
5	6.608	5.786	5.409	5.192	5.050	4.950	4.876	4.818
6	5.987	5.143	4.757	4.534	4.387	4.284	4.207	4.147
7	5.591	4.737	4.347	4.120	3.972	3.866	3.787	3.726
8	5.318	4.459	4.066	3.838	3.687	3.581	3.500	3.438
9	5.117	4.256	3.863	3.633	3.482	3.374	3.293	3.230
10	4.965	4.103	3.708	3.478	3.326	3.217	3.135	3.072
11	4.844	3.982	3.587	3.357	3.204	3.095	3.012	2.948
12	4.747	3.885	3.490	3.259	3.106	2.996	2.913	2.849
13	4.667	3.806	3.411	3.179	3.025	2.915	2.832	2.767
14	4.600	3.739	3.344	3.112	2.958	2.848	2.764	2.699
15	4.543	3.682	3.287	3.056	2.901	2.790	2.707	2.641
16	4.494	3.634	3.239	3.007	2.852	2.741	2.657	2.591
17	4.451	3.592	3.197	2.965	2.810	2.699	2.614	2.548
18	4.414	3.555	3.160	2.928	2.773	2.661	2.577	2.510
19	4.381	3.522	3.127	2.895	2.740	2.628	2.544	2.477
20	4.351	3.493	3.098	2.866	2.711	2.599	2.514	2.447
21	4.325	3.467	3.072	2.840	2.685	2.573	2.488	2.420
22	4.301	3.443	3.049	2.817	2.661	2.549	2.464	2.397
23	4.279	3.422	3.028	2.796	2.640	2.528	2.442	2.375
24	4.260	3.403	3.009	2.776	2.621	2.508	2.423	2.355
25	4.242	3.385	2.991	2.759	2.603	2.490	2.405	2.337
26	4.225	3.369	2.975	2.743	2.587	2.474	2.388	2.321
27	4.210	3.354	2.960	2.728	2.572	2.459	2.373	2.305
28	4.196	3.340	2.947	2.714	2.558	2.445	2.359	2.291
29	4.183	3.328	2.934	2.701	2.545	2.432	2.346	2.278
30	4.171	3.316	2.922	2.690	2.534	2.421	2.334	2.266
31	4.160	3.305	2.911	2.679	2.523	2.409	2.323	2.255
32	4.149	3.295	2.901	2.668	2.512	2.399	2.313	2.244
33	4.139	3.285	2.892	2.659	2.503	2.389	2.303	2.235
34	4.130	3.276	2.883	2.650	2.494	2.380	2.294	2.225
35	4.121	3.267	2.874	2.641	2.485	2.372	2.285	2.217
36	4.113	3.259	2.866	2.634	2.477	2.364	2.277	2.209
37	4.105	3.252	2.859	2.626	2.470	2.356	2.270	2.201
38	4.098	3.245	2.852	2.619	2.463	2.349	2.262	2.194
39	4.091	3.238	2.845	2.612	2.456	2.342	2.255	2.187
40	4.085	3.232	2.839	2.606	2.449	2.336	2.249	2.180
50	4.034	3.183	2.790	2.557	2.400	2.286	2.199	2.130
60	4.001	3.150	2.758	2.525	2.368	2.254	2.167	2.097
120	3.9301	3.072	2.6681	2.447	2.290	2.175	2.087	2.016

16
246.462
19.433
8.692
5.844
4.604
3.922
3.494
3.202
2.989
2.828
2.701
2.599
2.515
2.445
2.385
2.333
2.289
2.250
2.215
2.184
2.156
2.131
2.109
2.088
2.069
2.052
2.036
2.021
2.007
1.995
1.983
1.972
1.961
1.952
1.942
1.934
1.926
1.918
1.911
1.904
1.850
1.815
1.728

TABLE IX
F distribution ***(continued)***

γ_2 \ γ_1	17	18	19	20	21	22	23	24
1	246.917	247.322	247.685	248.012	248.308	248.577	248.824	249.051
2	19.437	19.440	19.443	19.446	19.448	19.450	19.452	19.454
3	8.683	8.675	8.667	8.660	8.654	8.648	8.643	8.639
4	5.832	5.821	5.811	5.803	5.795	5.787	5.781	5.774
5	4.590	4.579	4.568	4.558	4.549	4.541	4.534	4.527
6	3.908	3.896	3.884	3.874	3.865	3.856	3.849	3.841
7	3.480	3.467	3.455	3.445	3.435	3.426	3.418	3.411
8	3.187	3.173	3.161	3.150	3.140	3.131	3.123	3.115
9	2.974	2.960	2.948	2.936	2.926	2.917	2.908	2.900
10	2.812	2.798	2.785	2.774	2.764	2.754	2.745	2.737
11	2.685	2.671	2.658	2.646	2.636	2.626	2.617	2.609
12	2.583	2.568	2.555	2.544	2.533	2.523	2.514	2.505
13	2.499	2.484	2.471	2.459	2.448	2.438	2.429	2.420
14	2.428	2.413	2.400	2.388	2.377	2.367	2.357	2.349
15	2.368	2.353	2.340	2.328	2.316	2.306	2.297	2.288
16	2.317	2.302	2.288	2.276	2.264	2.254	2.244	2.235
17	2.272	2.257	2.243	2.230	2.219	2.208	2.199	2.190
18	2.233	2.217	2.203	2.191	2.179	2.168	2.159	2.150
19	2.198	2.182	2.168	2.156	2.144	2.133	2.123	2.114
20	2.167	2.151	2.137	2.124	2.112	2.102	2.092	2.082
21	2.139	2.123	2.109	2.096	2.084	2.073	2.063	2.054
22	2.114	2.098	2.084	2.071	2.059	2.048	2.038	2.028
23	2.091	2.075	2.061	2.048	2.036	2.025	2.014	2.005
24	2.070	2.054	2.040	2.027	2.015	2.003	1.993	1.984
25	2.051	2.035	2.021	2.007	1.995	1.984	1.974	1.964
26	2.034	2.018	2.003	1.990	1.978	1.966	1.956	1.946
27	2.018	2.002	1.987	1.974	1.961	1.950	1.940	1.930
28	2.003	1.987	1.972	1.959	1.946	1.935	1.924	1.915
29	1.989	1.973	1.958	1.945	1.932	1.921	1.910	1.901
30	1.976	1.960	1.945	1.932	1.919	1.908	1.897	1.887
31	1.965	1.948	1.933	1.920	1.907	1.896	1.885	1.875
32	1.953	1.937	1.922	1.908	1.896	1.884	1.873	1.864
33	1.943	1.926	1.911	1.898	1.885	1.873	1.863	1.853
34	1.933	1.917	1.902	1.888	1.875	1.863	1.853	1.843
35	1.924	1.907	1.892	1.878	1.866	1.854	1.843	1.833
36	1.915	1.899	1.883	1.870	1.857	1.845	1.834	1.824
37	1.907	1.890	1.875	1.861	1.848	1.837	1.826	1.816
38	1.899	1.883	1.867	1.853	1.841	1.829	1.818	1.808
39	1.892	1.875	1.860	1.846	1.833	1.821	1.810	1.800
40	1.885	1.868	1.853	1.839	1.826	1.814	1.803	1.793
50	1.831	1.814	1.798	1.784	1.771	1.759	1.748	1.737
60	1.796	1.778	1.763	1.748	1.735	1.722	1.711	1.700
120	1.709	1.690	1.674	1.659	1.645	1.632	1.620	1.608