Practice AP Statistics Exam
Saturday May 4, 2019
University of Delaware

Section I: Multiple Choice

Name: _____

School: _____

Instructions:

1. Choose the best answer to each multiple choice question.

2. Circle your choice on your exam form so you can score your exam later.

3. Record your choice using the "Record your MC answers here" link at the Registration site:

sites.udel.edu/crissing

DO NOT BEGIN UNTIL INSTRUCTED TO DO SO.

Formulas

(I) Descriptive Statistics

$$\bar{x} = \frac{\Sigma x_i}{n}$$

$$s_x = \sqrt{\frac{1}{n-1}\Sigma\left(x_i - \bar{x}\right)^2}$$

$$s_p = \sqrt{\frac{\left(n_1 - 1\right)s_1^2 + \left(n_2 - 1\right)s_2^2}{\left(n_1 - 1\right) + \left(n_2 - 1\right)}}$$

$$\hat{y} = b_0 + b_1 x$$

$$b_1 = \frac{\Sigma\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)}{\Sigma\left(x_i - \bar{x}\right)^2}$$

$$b_0 = \bar{y} - b_1\bar{x}$$

$$r = \frac{1}{n-1}\Sigma\left(\frac{x_i - \bar{x}}{s_x}\right)\left(\frac{y_i - \bar{y}}{s_y}\right)$$

$$b_1 = r\frac{s_y}{s_x}$$

$$s_{b_1} = \frac{\sqrt{\frac{\Sigma\left(y_i - \hat{y}_i\right)^2}{n-2}}}{\sqrt{\Sigma\left(x_i - \bar{x}\right)^2}}$$

(II) Probability

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

$$E(X) = \mu_x = \sum x_i p_i$$

$$\text{Var}(X) = \sigma_x^2 = \sum \left(x_i - \mu_x\right)^2 p_i$$

If X has a binomial distribution with parameters n and p, then:

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

$$\mu_x = np$$

$$\sigma_x = \sqrt{np(1-p)}$$

$$\mu_{\hat{p}} = p$$

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$$

If $\overline{x}$ is the mean of a random sample of size n from an infinite population with mean μ and standard deviation σ, then:

$$\mu_{\overline{x}} = \mu$$

$$\sigma_{\overline{x}} = \frac{\sigma}{\sqrt{n}}$$

(III) Inferential Statistics

Standardized test statistic: $\dfrac{\text{statistic} - \text{parameter}}{\text{standard deviation of statistic}}$

Confidence interval: statistic $\pm$ (critical value) $\bullet$ (standard deviation of statistic)

Single-Sample

Statistic	Standard Deviation of Statistic
Sample Mean	$\dfrac{\sigma}{\sqrt{n}}$
Sample Proportion	$\sqrt{\dfrac{p(1-p)}{n}}$

Two-Sample

Statistic	Standard Deviation of Statistic
Difference of sample means	$\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}$ Special case when $\sigma_1 = \sigma_2$ $\sigma\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}$
Difference of sample proportions	$\sqrt{\dfrac{p_1(1-p_1)}{n_1} + \dfrac{p_2(1-p_2)}{n_2}}$ Special case when $p_1 = p_2$ $\sqrt{p(1-p)}\sqrt{\dfrac{1}{n_1} + \dfrac{1}{n_2}}$

Chi-square test statistic $= \sum \dfrac{(\text{observed} - \text{expected})^2}{\text{expected}}$

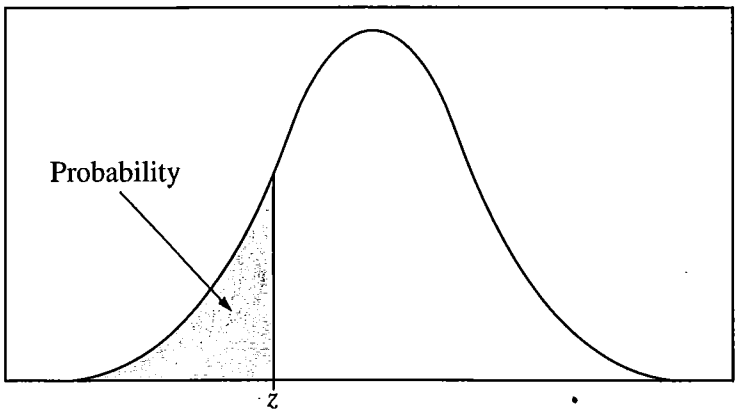


Table entry for z is the probability lying below z.

Table A Standard normal probabilities

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.4	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0003	.0002
−3.3	.0005	.0005	.0005	.0004	.0004	.0004	.0004	.0004	.0004	.0003
−3.2	.0007	.0007	.0006	.0006	.0006	.0006	.0006	.0005	.0005	.0005
−3.1	.0010	.0009	.0009	.0009	.0008	.0008	.0008	.0008	.0007	.0007
−3.0	.0013	.0013	.0013	.0012	.0012	.0011	.0011	.0011	.0010	.0010
−2.9	.0019	.0018	.0018	.0017	.0016	.0016	.0015	.0015	.0014	.0014
−2.8	.0026	.0025	.0024	.0023	.0023	.0022	.0021	.0021	.0020	.0019
−2.7	.0035	.0034	.0033	.0032	.0031	.0030	.0029	.0028	.0027	.0026
−2.6	.0047	.0045	.0044	.0043	.0041	.0040	.0039	.0038	.0037	.0036
−2.5	.0062	.0060	.0059	.0057	.0055	.0054	.0052	.0051	.0049	.0048
−2.4	.0082	.0080	.0078	.0075	.0073	.0071	.0069	.0068	.0066	.0064
−2.3	.0107	.0104	.0102	.0099	.0096	.0094	.0091	.0089	.0087	.0084
−2.2	.0139	.0136	.0132	.0129	.0125	.0122	.0119	.0116	.0113	.0110
−2.1	.0179	.0174	.0170	.0166	.0162	.0158	.0154	.0150	.0146	.0143
−2.0	.0228	.0222	.0217	.0212	.0207	.0202	.0197	.0192	.0188	.0183
−1.9	.0287	.0281	.0274	.0268	.0262	.0256	.0250	.0244	.0239	.0233
−1.8	.0359	.0351	.0344	.0336	.0329	.0322	.0314	.0307	.0301	.0294
−1.7	.0446	.0436	.0427	.0418	.0409	.0401	.0392	.0384	.0375	.0367
−1.6	.0548	.0537	.0526	.0516	.0505	.0495	.0485	.0475	.0465	.0455
−1.5	.0668	.0655	.0643	.0630	.0618	.0606	.0594	.0582	.0571	.0559
−1.4	.0808	.0793	.0778	.0764	.0749	.0735	.0721	.0708	.0694	.0681
−1.3	.0968	.0951	.0934	.0918	.0901	.0885	.0869	.0853	.0838	.0823
−1.2	.1151	.1131	.1112	.1093	.1075	.1056	.1038	.1020	.1003	.0985
−1.1	.1357	.1335	.1314	.1292	.1271	.1251	.1230	.1210	.1190	.1170
−1.0	.1587	.1562	.1539	.1515	.1492	.1469	.1446	.1423	.1401	.1379
−0.9	.1841	.1814	.1788	.1762	.1736	.1711	.1685	.1660	.1635	.1611
−0.8	.2119	.2090	.2061	.2033	.2005	.1977	.1949	.1922	.1894	.1867
−0.7	.2420	.2389	.2358	.2327	.2296	.2266	.2236	.2206	.2177	.2148
−0.6	.2743	.2709	.2676	.2643	.2611	.2578	.2546	.2514	.2483	.2451
−0.5	.3085	.3050	.3015	.2981	.2946	.2912	.2877	.2843	.2810	.2776
−0.4	.3446	.3409	.3372	.3336	.3300	.3264	.3228	.3192	.3156	.3121
−0.3	.3821	.3783	.3745	.3707	.3669	.3632	.3594	.3557	.3520	.3483
−0.2	.4207	.4168	.4129	.4090	.4052	.4013	.3974	.3936	.3897	.3859
−0.1	.4602	.4562	.4522	.4483	.4443	.4404	.4364	.4325	.4286	.4247
−0.0	.5000	.4960	.4920	.4880	.4840	.4801	.4761	.4721	.4681	.4641

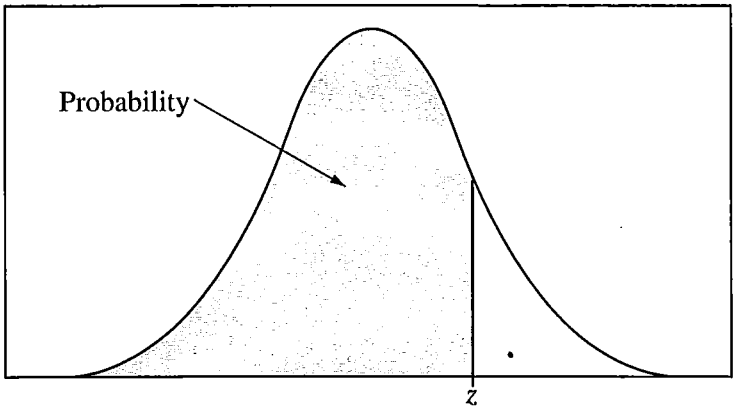


Table entry for z is the probability lying below z.

Table A (*Continued*)

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
0.0	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.1	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.2	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.3	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.4	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.5	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.6	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.7	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.8	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8078	.8106	.8133
0.9	.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

Table entry for p and C is the point t^* with probability p lying above it and probability C lying between $-t^*$ and t^*.

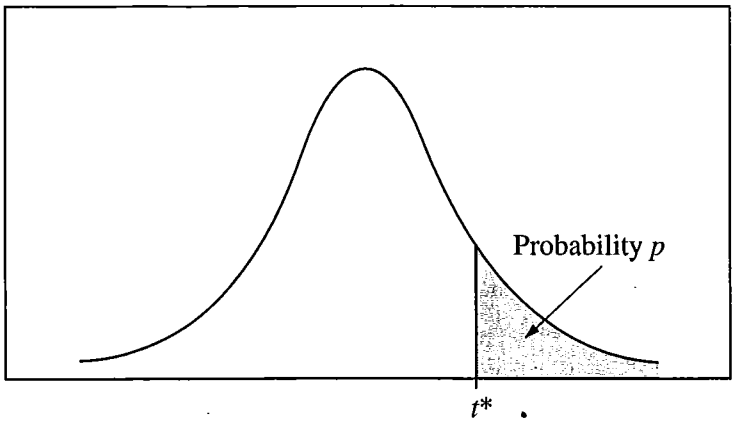


Table B t **distribution critical values**

	Tail probability p											
df	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.000	1.376	1.963	3.078	6.314	12.71	15.89	31.82	63.66	127.3	318.3	636.6
2	.816	1.061	1.386	1.886	2.920	4.303	4.849	6.965	9.925	14.09	22.33	31.60
3	.765	.978	1.250	1.638	2.353	3.182	3.482	4.541	5.841	7.453	10.21	12.92
4	.741	.941	1.190	1.533	2.132	2.776	2.999	3.747	4.604	5.598	7.173	8.610
5	.727	.920	1.156	1.476	2.015	2.571	2.757	3.365	4.032	4.773	5.893	6.869
6	.718	.906	1.134	1.440	1.943	2.447	2.612	3.143	3.707	4.317	5.208	5.959
7	.711	.896	1.119	1.415	1.895	2.365	2.517	2.998	3.499	4.029	4.785	5.408
8	.706	.889	1.108	1.397	1.860	2.306	2.449	2.896	3.355	3.833	4.501	5.041
9	.703	.883	1.100	1.383	1.833	2.262	2.398	2.821	3.250	3.690	4.297	4.781
10	.700	.879	1.093	1.372	1.812	2.228	2.359	2.764	3.169	3.581	4.144	4.587
11	.697	.876	1.088	1.363	1.796	2.201	2.328	2.718	3.106	3.497	4.025	4.437
12	.695	.873	1.083	1.356	1.782	2.179	2.303	2.681	3.055	3.428	3.930	4.318
13	.694	.870	1.079	1.350	1.771	2.160	2.282	2.650	3.012	3.372	3.852	4.221
14	.692	.868	1.076	1.345	1.761	2.145	2.264	2.624	2.977	3.326	3.787	4.140
15	.691	.866	1.074	1.341	1.753	2.131	2.249	2.602	2.947	3.286	3.733	4.073
16	.690	.865	1.071	1.337	1.746	2.120	2.235	2.583	2.921	3.252	3.686	4.015
17	.689	.863	1.069	1.333	1.740	2.110	2.224	2.567	2.898	3.222	3.646	3.965
18	.688	.862	1.067	1.330	1.734	2.101	2.214	2.552	2.878	3.197	3.611	3.922
19	.688	.861	1.066	1.328	1.729	2.093	2.205	2.539	2.861	3.174	3.579	3.883
20	.687	.860	1.064	1.325	1.725	2.086	2.197	2.528	2.845	3.153	3.552	3.850
21	.686	.859	1.063	1.323	1.721	2.080	2.189	2.518	2.831	3.135	3.527	3.819
22	.686	.858	1.061	1.321	1.717	2.074	2.183	2.508	2.819	3.119	3.505	3.792
23	.685	.858	1.060	1.319	1.714	2.069	2.177	2.500	2.807	3.104	3.485	3.768
24	.685	.857	1.059	1.318	1.711	2.064	2.172	2.492	2.797	3.091	3.467	3.745
25	.684	.856	1.058	1.316	1.708	2.060	2.167	2.485	2.787	3.078	3.450	3.725
26	.684	.856	1.058	1.315	1.706	2.056	2.162	2.479	2.779	3.067	3.435	3.707
27	.684	.855	1.057	1.314	1.703	2.052	2.158	2.473	2.771	3.057	3.421	3.690
28	.683	.855	1.056	1.313	1.701	2.048	2.154	2.467	2.763	3.047	3.408	3.674
29	.683	.854	1.055	1.311	1.699	2.045	2.150	2.462	2.756	3.038	3.396	3.659
30	.683	.854	1.055	1.310	1.697	2.042	2.147	2.457	2.750	3.030	3.385	3.646
40	.681	.851	1.050	1.303	1.684	2.021	2.123	2.423	2.704	2.971	3.307	3.551
50	.679	.849	1.047	1.299	1.676	2.009	2.109	2.403	2.678	2.937	3.261	3.496
60	.679	.848	1.045	1.296	1.671	2.000	2.099	2.390	2.660	2.915	3.232	3.460
80	.678	.846	1.043	1.292	1.664	1.990	2.088	2.374	2.639	2.887	3.195	3.416
100	.677	.845	1.042	1.290	1.660	1.984	2.081	2.364	2.626	2.871	3.174	3.390
1000	.675	.842	1.037	1.282	1.646	1.962	2.056	2.330	2.581	2.813	3.098	3.300
∞	.674	.841	1.036	1.282	1.645	1.960	2.054	2.326	2.576	2.807	3.091	3.291
	50%	60%	70%	80%	90%	95%	96%	98%	99%	99.5%	99.8%	99.9%
	Confidence level C											

Table entry for p is the point (χ^2) with probability p lying above it.

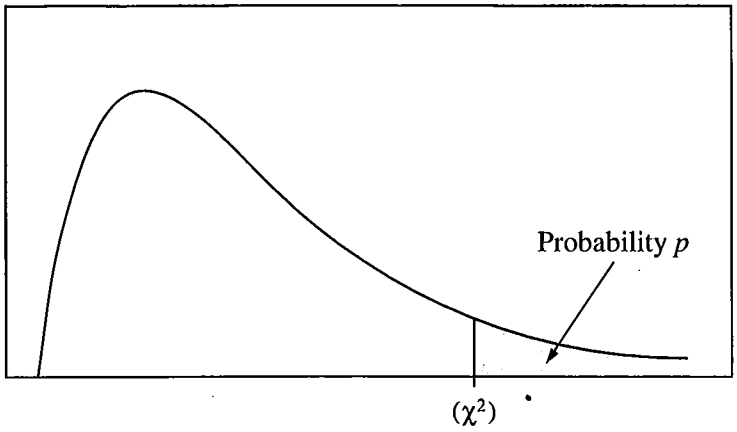


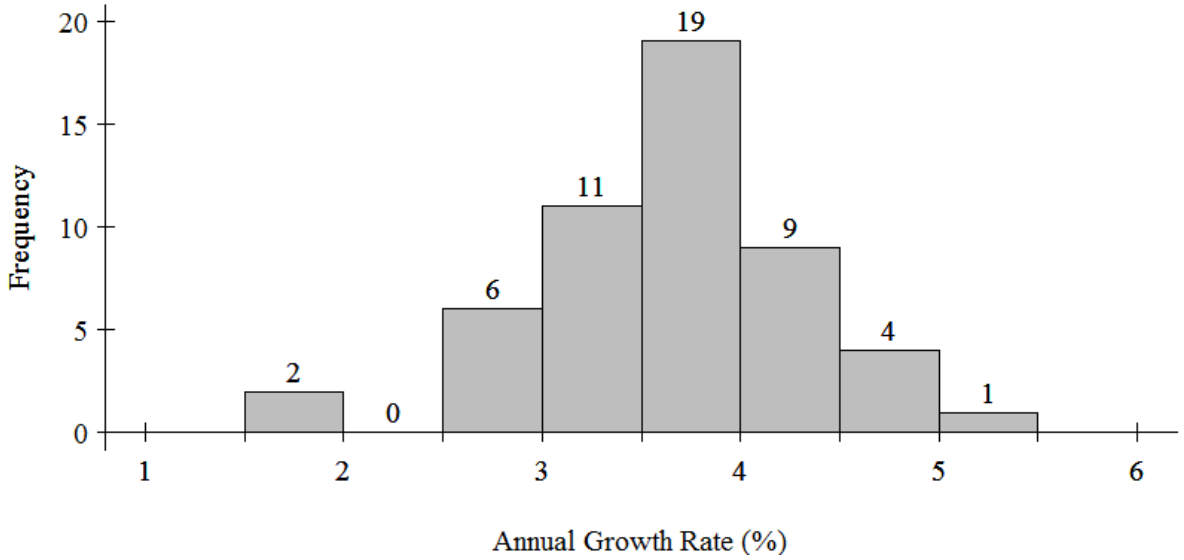
Table C χ^2 **critical values**

	Tail probability p											
df	.25	.20	.15	.10	.05	.025	.02	.01	.005	.0025	.001	.0005
1	1.32	1.64	2.07	2.71	3.84	5.02	5.41	6.63	7.88	9.14	10.83	12.12
2	2.77	3.22	3.79	4.61	5.99	7.38	7.82	9.21	10.60	11.98	13.82	15.20
3	4.11	4.64	5.32	6.25	7.81	9.35	9.84	11.34	12.84	14.32	16.27	17.73
4	5.39	5.99	6.74	7.78	9.49	11.14	11.67	13.28	14.86	16.42	18.47	20.00
5	6.63	7.29	8.12	9.24	11.07	12.83	13.39	15.09	16.75	18.39	20.51	22.11
6	7.84	8.56	9.45	10.64	12.59	14.45	15.03	16.81	18.55	20.25	22.46	24.10
7	9.04	9.80	10.75	12.02	14.07	16.01	16.62	18.48	20.28	22.04	24.32	26.02
8	10.22	11.03	12.03	13.36	15.51	17.53	18.17	20.09	21.95	23.77	26.12	27.87
9	11.39	12.24	13.29	14.68	16.92	19.02	19.68	21.67	23.59	25.46	27.88	29.67
10	12.55	13.44	14.53	15.99	18.31	20.48	21.16	23.21	25.19	27.11	29.59	31.42
11	13.70	14.63	15.77	17.28	19.68	21.92	22.62	24.72	26.76	28.73	31.26	33.14
12	14.85	15.81	16.99	18.55	21.03	23.34	24.05	26.22	28.30	30.32	32.91	34.82
13	15.98	16.98	18.20	19.81	22.36	24.74	25.47	27.69	29.82	31.88	34.53	36.48
14	17.12	18.15	19.41	21.06	23.68	26.12	26.87	29.14	31.32	33.43	36.12	38.11
15	18.25	19.31	20.60	22.31	25.00	27.49	28.26	30.58	32.80	34.95	37.70	39.72
16	19.37	20.47	21.79	23.54	26.30	28.85	29.63	32.00	34.27	36.46	39.25	41.31
17	20.49	21.61	22.98	24.77	27.59	30.19	31.00	33.41	35.72	37.95	40.79	42.88
18	21.60	22.76	24.16	25.99	28.87	31.53	32.35	34.81	37.16	39.42	42.31	44.43
19	22.72	23.90	25.33	27.20	30.14	32.85	33.69	36.19	38.58	40.88	43.82	45.97
20	23.83	25.04	26.50	28.41	31.41	34.17	35.02	37.57	40.00	42.34	45.31	47.50
21	24.93	26.17	27.66	29.62	32.67	35.48	36.34	38.93	41.40	43.78	46.80	49.01
22	26.04	27.30	28.82	30.81	33.92	36.78	37.66	40.29	42.80	45.20	48.27	50.51
23	27.14	28.43	29.98	32.01	35.17	38.08	38.97	41.64	44.18	46.62	49.73	52.00
24	28.24	29.55	31.13	33.20	36.42	39.36	40.27	42.98	45.56	48.03	51.18	53.48
25	29.34	30.68	32.28	34.38	37.65	40.65	41.57	44.31	46.93	49.44	52.62	54.95
26	30.43	31.79	33.43	35.56	38.89	41.92	42.86	45.64	48.29	50.83	54.05	56.41
27	31.53	32.91	34.57	36.74	40.11	43.19	44.14	46.96	49.64	52.22	55.48	57.86
28	32.62	34.03	35.71	37.92	41.34	44.46	45.42	48.28	50.99	53.59	56.89	59.30
29	33.71	35.14	36.85	39.09	42.56	45.72	46.69	49.59	52.34	54.97	58.30	60.73
30	34.80	36.25	37.99	40.26	43.77	46.98	47.96	50.89	53.67	56.33	59.70	62.16
40	45.62	47.27	49.24	51.81	55.76	59.34	60.44	63.69	66.77	69.70	73.40	76.09
50	56.33	58.16	60.35	63.17	67.50	71.42	72.61	76.15	79.49	82.66	86.66	89.56
60	66.98	68.97	71.34	74.40	79.08	83.30	84.58	88.38	91.95	95.34	99.61	102.7
80	88.13	90.41	93.11	96.58	101.9	106.6	108.1	112.3	116.3	120.1	124.8	128.3
100	109.1	111.7	114.7	118.5	124.3	129.6	131.1	135.8	140.2	144.3	149.4	153.2

AP® STATISTICS
SECTION I
Time --- 1 hour and 30 minutes
Number of questions --- 40
Percent of total grade --- 50

Directions: Solve each of the following problems, using the available space for scratch work. Decide which is the best of the choices given and fill in the corresponding box on the answer sheet. Do not spend too much time on any one problem.

1. The histogram below shows the distribution of annual growth rate in travel spending for 52 companies in the United States.



After creating this graph, it was discovered that the values for the companies with the 20 largest reported annual growth rates were miscalculated. In fact, each of these top 20 companies should have a higher value of annual growth rate than was previously reported. Which of the following statements must be true?

(A) If you were to make a graph with the corrected values, the graph would be skewed to the left.

(B) If you were to make a graph with the corrected values, the graph would be symmetric.

(C) If you were to calculate the median with the corrected values, the new median would be greater than the old median.

(D) If you were to calculate the interquartile range (IQR) with the corrected values, the new IQR would be the same as the old IQR.

(E) If you were to find the value of the third quartile (Q3) with the corrected values, the new Q3 would be greater than the old Q3.

2. The council leaders of a small town would like to determine the average number of people living in each household. To accomplish this, they send a questionnaire to each residence listed in the town address records, asking the head of the household to record the number of people living in the dwelling. They follow this up with visits to those households that have not responded to the questionnaire until the data collection process is complete. Which of the following is the best description of the method for data collection used by the council leaders?

   (A) A convenience sample

   (B) A simple random sample

   (C) A stratified random sample

   (D) A cluster random sample

   (E) A census

---

3. A biologist is interested in predicting the percentage increase in lung volume when inhaling (y) for a certain species of bird from the percentage of carbon dioxide in the atmosphere (x). Data collected from a random sample of 20 birds of this species were used to create the least-squares regression equation $\hat{y} = 400 - 0.08x$. Which of the following best describes the meaning of the slope of the least-squares regression line?

   (A) The percentage increase in lung volume when inhaling increases by 0.08 percent, on average, for every 1 percent increase in the carbon dioxide in the atmosphere.

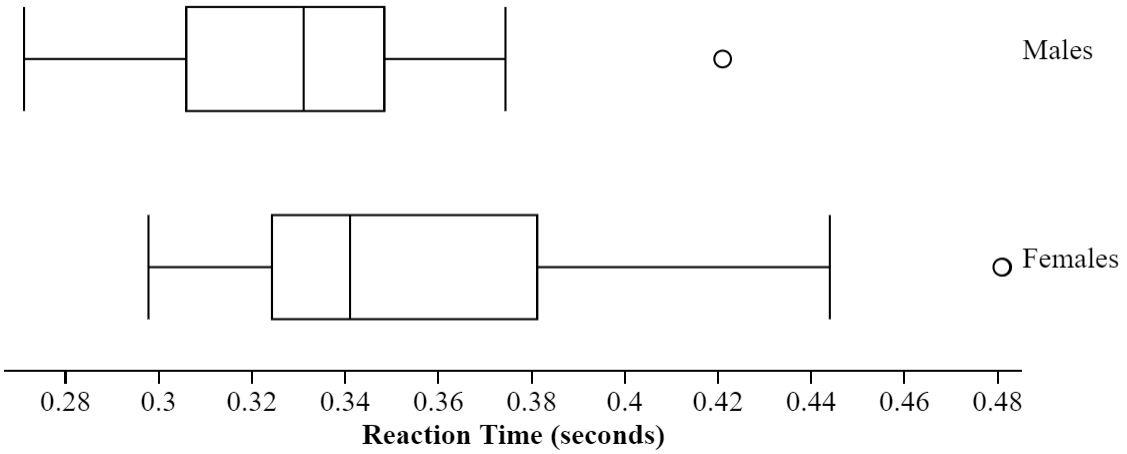
   (B) The percentage of carbon dioxide in the atmosphere increases by 0.08 percent, on average, for every 1 percent increase in lung volume when inhaling.

   (C) The percentage increase in lung volume when inhaling decreases by 0.08 percent, on average, for every 1 percent increase in the carbon dioxide in the atmosphere.

   (D) The percentage of carbon dioxide in the atmosphere increases by 0.08 percent, on average, for every 1 percent increase in lung volume when inhaling.

   (E) Approximately 8% of the variability in the percentage increase in lung volume when inhaling is explained by its linear relationship with the percentage of carbon dioxide in the atmosphere.

4. A student was interested in investigating whether female or male students at her school tend to have quicker reaction times. The students in her study played an online game that measured reaction time by how quickly you press stop after an orange sign appears. The timing of the appearance of the orange signs is random. The reaction times (in seconds) for the male and female students at her school that volunteered to take part in the study are displayed in the boxplots below.



Based on the boxplots, which of the following statements must be true?

(A) At least 75% of the female reaction times are greater than the maximum reaction time for the males.

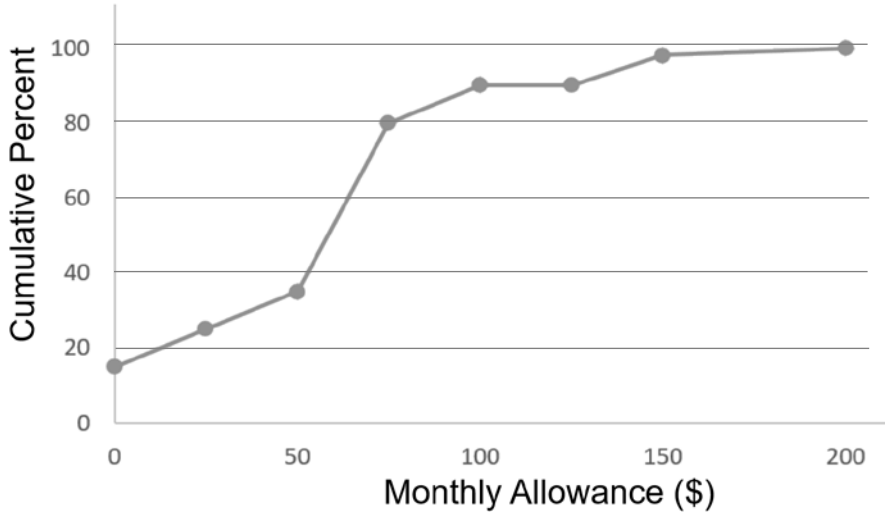
(B) If the highest reaction time for the males was removed, the distribution of male times would be approximately normally distributed.

(C) There is more variability in the highest 25% of reaction times for the female students than in the highest 25% of times for the male students.

(D) More female students took part in the study than male students.

(E) The mean reaction time for the male students is greater than the mean reaction time for the female students

5. The monthly allowance in dollars for a random sample of high school students is displayed in the cumulative frequency graph below.



Based on this graphical display, which of the following statements is justified?

(A) A greater percentage of these students have monthly allowance between $50 and $75 than between $175 and $200.

(B) The median monthly allowance for these students is about $100.

(C) A greater percentage of these students have an allowance above $100 than below $100.

(D) The minimum monthly allowance for these students is greater than $10.

(E) The graph of the cumulative percent of monthly allowance for these students is skewed to the right.

6. Doctors are now considering prescribing social activities as a way to increase the overall health of their patients. A random sample of adult men with heart disease were prescribed two sessions a week consisting of 30 minute group discussions on lifestyle changes followed by 45 minute golf lessons. As an indicator of overall health, each subjects' grip strength was measured at the start and end of the six-week study, with higher grip strength indicating better health. Let μ_B represent the population mean grip strength before participating in prescribed social activities, μ_A represent the population mean grip strength after participating in prescribed social activities, and μ_D represent the population mean difference in grip strength (before – after). Which of the following would be the null and alternative hypotheses for this study?

   (A) $H_0 : \mu_B = \mu_A$ and $H_a : \mu_B > \mu_A$

   (B) $H_0 : \mu_B = \mu_A$ and $H_a : \mu_B \neq \mu_A$

   (C) $H_0 : \mu_D = \mu_B - \mu_A$ and $H_a : \mu_D < \mu_B - \mu_A$

   (D) $H_0 : \mu_D > 0$ and $H_a : \mu_D = 0$

   (E) $H_0 : \mu_D = 0$ and $H_a : \mu_D < 0$

---

7. In 2013, Ace Metrix gathered a random sample of 1200 television ads and found that 180 (or 15%) featured a celebrity. They created a confidence interval to estimate the proportion of all ads that feature a celebrity to within 2.7%. Based on this interval, which of the following must be true?

   (A) Approximately 2.7% of all television ads were sampled.

   (B) With 97.3% confidence, 15% of all television ads feature a celebrity.

   (C) It is plausible that 19% of all television ads feature a celebrity.

   (D) It is plausible that 13% of all television ads feature a celebrity.

   (E) Between 12.3% and 17.7% of all celebrities have been featured in television ads.

8. As a reward for completing a reading challenge, two elementary school students are allowed to select one ice cream bar from a box containing 12 bars. The box contains four chocolate bars (C), four strawberry bars (S), and four vanilla bars (V). Suppose the two students select their ice cream bars at random. The probability for each possible selection is given in the table below.

Bars Selected	CC	CV	CS	SC	SS	SV	VC	VS	VV
Probability	$\frac{1}{11}$	$\frac{4}{33}$	$\frac{4}{33}$	$\frac{4}{33}$	$\frac{1}{11}$	$\frac{4}{33}$	$\frac{4}{33}$	$\frac{4}{33}$	$\frac{1}{11}$

Find the expected number of vanilla bars selected.

(A) $\frac{1}{3}$ bars

(B) $\frac{2}{3}$ bars

(C) $\frac{6}{11}$ bars

(D) $\frac{19}{13}$ bars

(E) 1 bar

9. In an effort to increase sales, a company is considering a new training program for its sales department. An advertisement claims that the training program leads to a mean increase of \$5000 in monthly sales. To study the program's effectiveness, the company selects 20 sales representatives, randomly assigning 10 to the training program and 10 to serve as control. On average, the representatives who participated in the training brought in \$8231 more in monthly sales than the control, and the 95% confidence interval for the difference in means was (2799, 13663) dollars. Does this confidence interval provide convincing statistical evidence that the mean increase in sales is different from the advertised value?

(A) No, because \$0 is outside the interval the results are not statistically significant.

(B) No, because \$5000 is included in the interval the results are not statistically significant.

(C) Yes, because \$0 is outside the interval the results are statistically significant.

(D) Yes, because the representatives were randomly assigned and \$8231 is different from \$5000.

(E) Yes, because \$8231 is inside the confidence interval the results are accurate.

---

10. In a very large shipment of electronic devices, 5% are defective. Let X represent the number of defective devices in a random sample of 20 devices from this shipment. Which of the following expressions gives $P(X > 2)$?

(A) $(0.05)^2(0.95)^{18}$

(B) $\sum_{k=2}^{20} \binom{20}{k} (0.05)^k (0.95)^{20-k}$

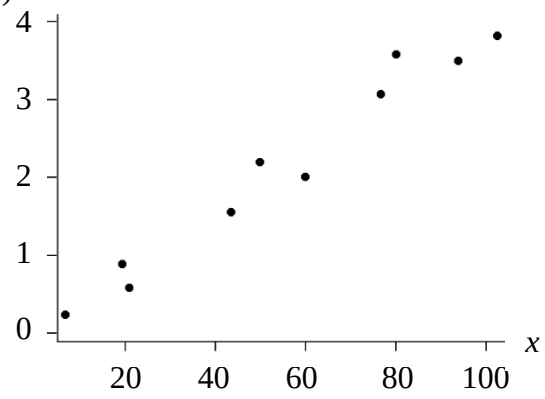
(C) $\sum_{k=3}^{20} \binom{20}{k} (0.05)^k (0.95)^{20-k}$

(D) $\sum_{k=3}^{20} \binom{20}{k} (0.95)^k (0.05)^{20-k}$

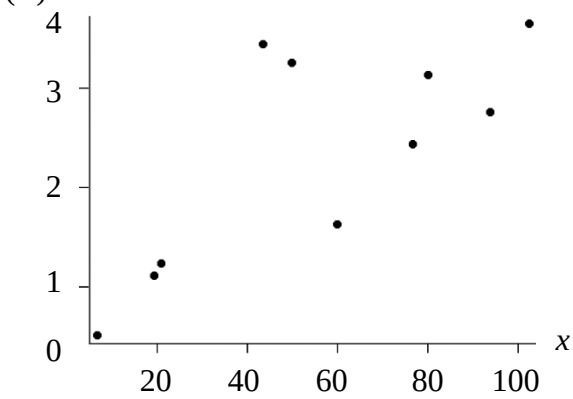
(E) $1 - \sum_{k=3}^{20} \binom{20}{k} (0.05)^k (0.95)^{20-k}$

11. Which of the following scatterplots is best described as displaying a strong, positive linear relationship between x and y?

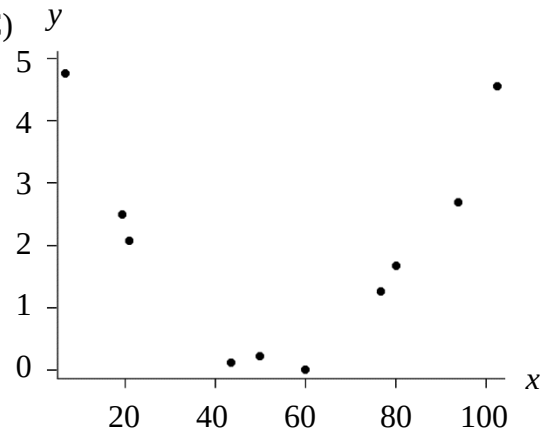
(A)



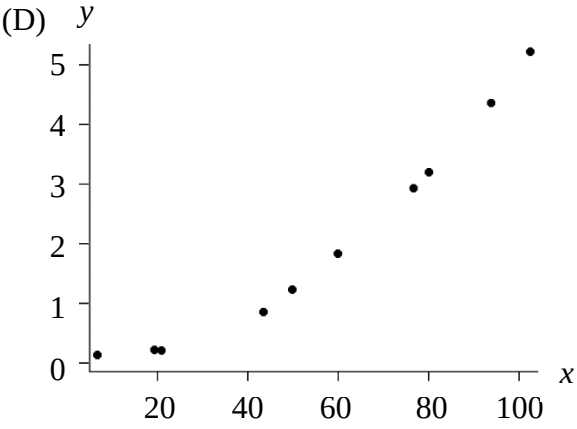
(B)



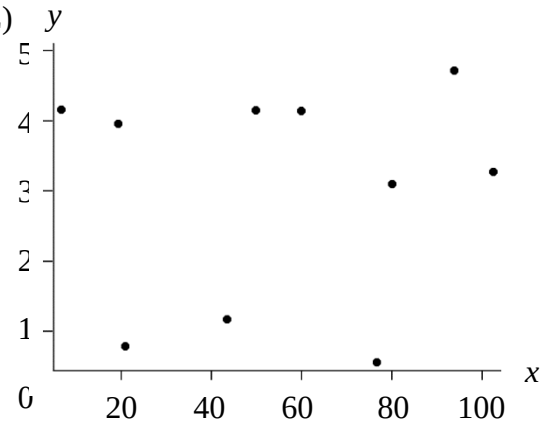
(C)



(D)



(E)



12. The female life expectancy and the male life expectancy (in years) for 30 industrial countries were recorded. The data are shown in the back-to-back stemplot below.

Female Life Expectancy		Male Life Expectancy
	68	6 9
	69	
	70	1 8
	71	
	72	9
8	73	0
	74	9
	75	1 2 5 6 8
9	76	2 2 7 7 8 9
9 9	77	1 2 4 5 6 7
	78	0 4 5 6 7
4 1	79	2
4 2	80	
9 8 8 7 7 6 6 4 1	81	
8 7 5 3 3 2	82	
9 9 8 3 2 1	83	
	84	
5	85	

68|6 represents 68.6 years

Based on the stemplot, which of the following statements is true?

(A) The median life expectancy and the range of life expectancies are both less for females than for males in these countries.

(B) The median life expectancy and the range of life expectancies are both greater for females than for males in these countries.

(C) The median life expectancy is the same for females and males in these countries, and the range of life expectancies is the same for females and for males in these countries.

(D) The median life expectancy is less for females than for males in these countries, and the range of life expectancies is greater for females than for males in these countries.

(E) The median life expectancy is greater for females than for males in these countries, and the range of life expectancies is less for females than for males in these countries.

13. A medical researcher conducted a study to investigate whether exercising can lower cholesterol. From a sample of 150 adult volunteers with high cholesterol, 75 were randomly assigned to pursue an exercise program that includes daily jogging and strength training. The remaining 75 were asked to refrain from vigorous activity. At the end of six months, the people in the exercise group substantially reduced their cholesterol, while the people in the no-exercise group did not experience a reduction. Which of the following is the most appropriate conclusion?

(A) There is evidence that exercise causes a reduction in cholesterol, and the conclusion can be generalized to all adults.

(B) There is evidence that exercise causes a reduction in cholesterol, and the conclusion can be generalized to all adults who exercise.

(C) There is evidence that exercise causes a reduction in cholesterol, and the conclusion can be generalized to adults similar to these study volunteers with high cholesterol.

(D) Although cause-and-effect cannot be established, there is an association between exercising and a reduction in cholesterol for the population of all adults.

(E) Although cause-and-effect cannot be established, there is an association between exercising and a reduction in cholesterol for the population of all adults who exercise.

---

14. A two-sided *t*-test of the null hypothesis $H_0 : \mu = 40$ is conducted. If a 95 percent *t*-interval constructed from the same sample data does not contain the value of 40, which of the following can be concluded about the test at a significance level $\alpha = 0.05$?

(A) The *p*-value is less than 0.05, and H_0 should be rejected.

(B) The *p*-value is less than 0.05, and H_0 should not be rejected.

(C) The *p*-value is greater than 0.05, and H_0 should be rejected.

(D) The *p*-value is greater than 0.05, and H_0 should not be rejected.

(E) There is not enough information given to make a decision about H_0.

15. In the weeks leading up to the election of a controversial public figure to political office, a polling agency conducted in-person interviews with 1000 randomly selected eligible voters. Only 35 percent of these 1000 voters said they planned to vote for the controversial candidate. However, on election day, when voters can make their selections anonymously and without having to reveal who they voted for to another person, the controversial candidate received the majority of the vote. Which of the following types of bias is most likely to have occurred in the polling agency's survey?

(A) Nonresponse bias

(B) Sampling bias

(C) Selection bias

(D) Response bias

(E) Undercoverage bias

---

16. A news station would like to conduct an exit poll to determine the likelihood that a highly debated amendment will receive enough support to pass. The news station plans to construct a 90 percent confidence interval to estimate the proportion of voters supporting the amendment. Which of the following expressions would give the smallest sample size that will result in a margin of error of no more than 3 percentage points?

(A) $(0.5)\left(\frac{1.645}{0.03}\right)^2$

(B) $(0.5)^2\left(\frac{1.645}{0.03}\right)^2$

(C) $(0.5)^2\left(\frac{1.645}{0.06}\right)^2$

(D) $(0.5)^2\left(\frac{1.96}{0.03}\right)^2$

(E) $(0.5)^2\left(\frac{1.96}{0.06}\right)$

17. If P(A) = 0.40, P(B) = 0.75, and P(A and B) = 0.30, then what is true about the events A and B?

(A) Event A and event B are mutually exclusive and are independent.

(B) Event A and event B are mutually exclusive and are not independent.

(C) Event A and event B are not mutually exclusive and are independent.

(D) Event A and event B are not mutually exclusive and are not independent.

(E) Event A and event B are not mutually exclusive, and independence cannot be determined with the information given.

---

18. According to a popular wedding website, the average amount spent on flowers for a wedding in the U.S. is $750. To investigate whether the mean amount spent is actually less than the reported average, the mean and standard deviation of the amount spent on flowers from a random sample of 30 recent weddings in the U.S. was obtained. A one-sample t-test resulted in a p-value of 0.065. What conclusion should be made at the 0.05 level of significance?

(A) There is convincing statistical evidence to suggest the mean amount spent on flowers for weddings in the U.S. is less than $750.

(B) There is convincing statistical evidence to suggest the mean amount spent on flowers for weddings in the U.S. is $750.

(C) There is convincing statistical evidence to suggest the mean amount spent on flowers for weddings in the U.S. is greater than $750.

(D) There is not convincing statistical evidence to suggest the mean amount spent on flowers for weddings in the U.S. is less than $750.

(E) There is not convincing statistical evidence to suggest the mean amount spent on flowers for weddings in the U.S. is $750.

19. A forester measured the trees in a large woods up for sale. He concluded that the diameter of the trees was normally distributed with mean 12.4 inches and standard deviation 4.7 inches. Which of the following equations can be used to find the diameter x which is at the 90^{th} percentile?

(A) $0.90 = \frac{x - 12.4}{(4.7)^2}$

(B) $1.28 = \frac{x - 12.4}{(4.7)^2}$

(C) $0.90 = \frac{x - 12.4}{4.7}$

(D) $1.28 = \frac{x - 12.4}{4.7}$

(E) $1.645 = \frac{x - 12.4}{4.7}$

---

20. A sports physiologist is investigating whether water or Gatorade provides the best hydration for athletes after training. A sample of 50 athletes between the ages of 18 and 30 was selected. On two consecutive days, each athlete ran 5 miles at a steady rate on a treadmill in a 75-degree room. On one of these days, 32 ounces of water was consumed after running. On the other day, 32 ounces of Gatorade was consumed after running. For each athlete, which beverage was consumed first was randomized. Hydration one hour after beverage consumption was measured, and results for the two beverages were compared. Which of the following is the best description of this study?

(A) An experiment using a completely randomized design

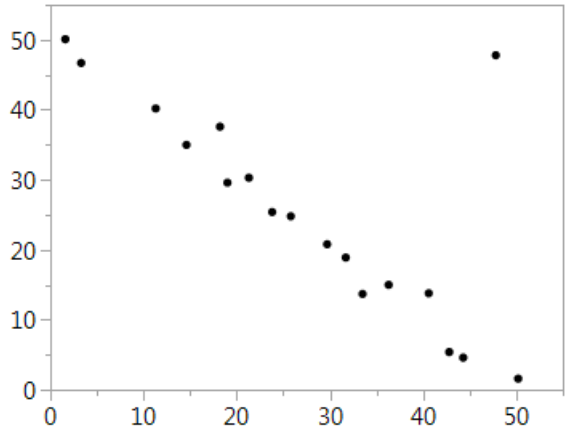
(B) An experiment using a matched-pairs design

(C) An observational study using a simple random sample

(D) An observational study using a cluster sample

(E) An observational study using a stratified sample

21. Consider the following scatterplot.



Which of the following is the most likely value for the correlation coefficient of this data set?

(A) -0.99

(B) -0.74

(C) 0.02

(D) 0.74

(E) 0.99

---

22. A recent report from the Pew Research Center indicates that even teens think they have a problem being too attached to their phones. A total of 365 teenage girls and 378 teenage boys were surveyed. Of the teenage girls surveyed, 171 reported feeling too attached to their phone. Of the teenage boys surveyed, 132 reported feeling too attached to their phone. Is there convincing statistical evidence of a difference between the two population proportions at the significance level of 0.01?

(A) Yes, because the sample proportions are different.

(B) Yes, because the *p*-value is less than 0.01.

(C) Yes, because the *p*-value is greater than 0.01.

(D) No, because the *p*-value is less than 0.01.

(E) No, because the *p*-value is greater than 0.01.

23. The distribution of heights of the players on a university football team in Arizona is approximately normally distributed with a mean of 74.9 inches and standard deviation of 2.0 inches. Robert plays for this Arizona team. His best friend Peter plays for a university football team in Florida. Both players are 73.2 inches tall and both player heights have the same standardized score (*z*-score) for their respective teams. If the Florida distribution is also approximately normal, but with a standard deviation of 2.9 inches, what is the mean height of a football player on this Florida team?

(A) 70.7 inches

(B) 72.0 inches

(C) 74.4 inches

(D) 75.7 inches

(E) 77.8 inches

---

24. Jake keeps a log of the distance he covers, in miles, on each of his daily running workouts. Over the past twelve months, he has recorded a large number of workouts. Jake randomly selected 10 workouts from his log and found a mean distance of 4.81 miles with a standard deviation of 2.05 miles. Assuming all conditions for inference are met, which of the following gives a 90% confidence interval to estimate the mean number of miles he ran in the past twelve months?

(A) $4.81 \pm 1.645(0.648)$

(B) $4.81 \pm 1.81(0.648)$

(C) $4.81 \pm 1.81(2.05)$

(D) $4.81 \pm 1.83(0.648)$

(E) $4.81 \pm 1.83(2.05)$

25. Scores on a standardized math placement test are normally distributed with mean 130 points and standard deviation 15 points. Which of the following is closest to the percent of test takers who score between 100 points and 145 points?

(A) 18%

(B) 32%

(C) 68%

(D) 82%

(E) 95%

---

26. The city of Pine Bluffs is studying the relationship between the number of police officers employed by the city and the number of crimes committed. The city council surveyed a random sample of other cities similar to Pine Bluffs to gather data on the number of crimes committed during a specified time period and the number of police officers the city employed. The results of an analysis of the data are shown below.

```
The regression equation is
Number of Crimes = 29.388 - 0.9596 Police

Predictor     Coef  SE Coef      T   P-Value
Constant    29.388    4.143   7.09     0.000
Police     -0.9596   0.2174  -4.41     0.004
```

What is the null hypothesis corresponding to the *p*-value of 0.004?

(A) Changing the number of police affects the number of crimes for all cities like Pine Bluffs.

(B) There is a linear relationship between the number of police and the number of crimes for all cities like Pine Bluffs.

(C) There is no linear relationship between the number of police and the number of crimes for all cities like Pine Bluffs.

(D) There is no relationship between the number of police and the number of crimes for all cities like Pine Bluffs.

(E) There is no difference between the mean number of police and the mean number of crimes for all cities like Pine Bluffs.

27. The city of Edmonton, Canada recently completed an audit of abandoned bank accounts. An account is considered abandoned if it has not been accessed in the last 20 years. A random sample of 50 such accounts was collected and a 95% confidence interval for the true mean abandoned account balance was found to be (267.43, 1455.00) dollars. Which of the following statements is a correct interpretation of the 95 percent confidence level?

(A) If many such samples of size 50 were taken, we would expect 95% of the intervals constructed to contain the mean balance of all abandoned bank accounts in Edmonton.

(B) There is a 95% chance that the constructed confidence interval contains the mean balance of all abandoned bank accounts in Edmonton.

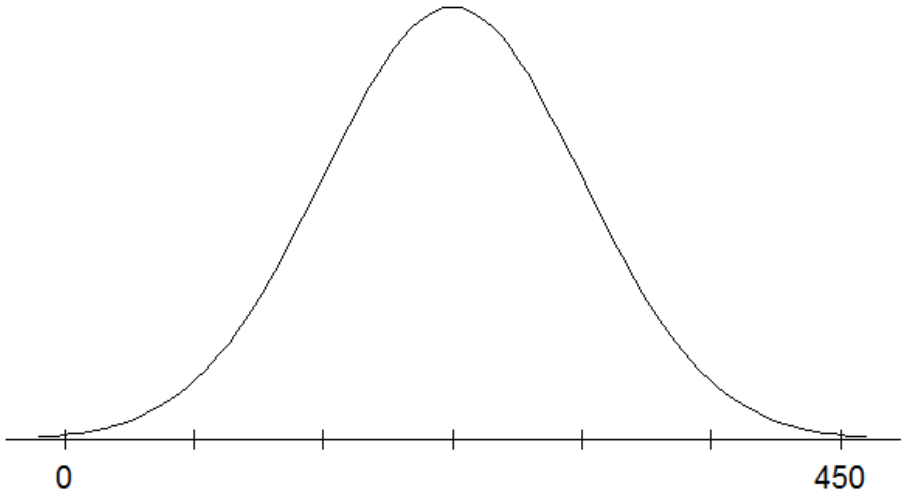
(C) If another random sample of 50 abandoned bank accounts were taken, we would be 95% confident that the sample mean balance would fall between $267.43 and $1455.00.

(D) If many such samples of 50 abandoned bank accounts were taken, we would expect 95% of the intervals constructed to contain the sample mean balance.

(E) There is a 95% chance that the sample mean balance of these 50 abandoned bank accounts in Edmonton lies between $267.43 and $1455.00.

---

28. Consider the following normal curve.



Which of the following choices is closest to the standard deviation of this normal distribution?

(A) 25

(B) 50

(C) 75

(D) 100

(E) 225

29. An administrator at a large high school was interested in the types of school-sponsored extracurricular activities in which students participated: only academic clubs, only intramural sports clubs, both type of clubs, or neither. The responses from a survey of 500 students selected at random are summarized in the frequency table below, shown by class standing.

	1st year	2nd year	3rd year	4th year	Total
Only Academics	30	40	60	60	190
Only Sports	35	30	20	20	105
Both	25	15	5	5	50
Neither	70	55	25	5	155
Total	160	140	110	90	500

If the school assumes extracurricular participation is independent of class standing, which of the following expressions could be used to calculate the expected number of respondents who do not participate in any academic or intramural activities during their 4th year at the school?

(A) $\frac{(5)(155)}{500}$

(B) $\frac{(90)(155)}{500}$

(C) $\frac{5}{155}$

(D) $\frac{5}{90}$

(E) $\frac{90}{155}$

30. Two different medications are to be compared in the treatment of depression among elderly patients in an assisted living facility. Each patient has his or her own room, but for safety concerns some are restricted to a phone only for in-house calls and some are allowed a phone that can call outside the facility. Which of the following would result in a randomized block design with blocks based on type of phone?

(A) Within patients that have the same type of phone, randomly assign half to each medication.

(B) Within patients taking the same type of medication, randomly assign half to each type of phone.

(C) Randomly assign all patients with phones that can call outside the facility to take one medication and have all patients with phones only for in-house calls take the other medication.

(D) Randomly assign all patients on one medication to phones that can call outside the facility and all patients on the other medication to phones only for in-house calls.

(E) Allow patients to choose which medication they prefer, and then compare the two medications within patients that have the same type of phone.

---

31. A blood bank is conducting a study on how much time is required of their patients for blood platelet donations. Based on past records, time spent in the waiting room has a mean of 5 minutes and a standard deviation of 1 minute, time spent with a nurse who reviews patients' medical records and measures their vital signs has a mean of 10 minutes and a standard deviation of 3 minutes, and the actual donation time has a mean of 130 minutes and a standard deviation of 8 minutes. Assuming these three times are independent and can all be modeled using a normal distribution, approximately what proportion of donors finish the entire donation process in under 2.5 hours (150 minutes)?

(A) 0.66

(B) 0.72

(C) 0.84

(D) 0.89

(E) 0.99

32. In June 2016, residents of the United Kingdom (UK) voted to leave the European Union (EU). However, 48.1% of the UK population voted to remain in the EU. In a random sample of 4399 eligible voters in the city of London, 3312 favor remaining in the EU. Which of the following expressions gives the test statistic for the appropriate test to determine if the proportion of 'remainers' in the city of London differs significantly from the proportion of the UK population who voted to remain in the EU?

(A) $z = \dfrac{0.753 - 0.481}{\sqrt{\dfrac{(0.50)(0.50)}{4399}}}$

(B) $z = \dfrac{0.753 - 0.481}{\sqrt{\dfrac{(0.481)(0.519)}{4399}}}$

(C) $z = \dfrac{0.753 - 0.481}{\sqrt{(0.481)(0.519)\left(\dfrac{1}{3312} + \dfrac{1}{1087}\right)}}$

(D) $z = \dfrac{0.481 - 0.753}{\sqrt{\dfrac{(0.753)(0.247)}{4399}}}$

(E) $z = \dfrac{0.753 - 0.481}{\sqrt{(0.50)(0.50)\left(\dfrac{1}{3312} + \dfrac{1}{1087}\right)}}$

33. The average daily high temperature in degrees Celsius (C) during the month of July was recorded for the 51 largest cities in the United States. Summary measures are shown below.

Mean	Standard Deviation	Median	IQR	Range
30.84	3.78	31	5	22

Which of the following represents the median and interquartile range (IQR) of these average daily high temperatures in degrees Fahrenheit (F), where $F = \left(\frac{9}{5}\right)C + 32$?

(A) Median = 52.8°F, IQR = 9°F

(B) Median = 52.8°F, IQR = 41°F

(C) Median = 87.8°F, IQR = 5°F

(D) Median = 87.8°F, IQR = 9°F

(E) Median = 87.8°F, IQR = 41°F

---

34. In a large introductory statistics course at one university, 26% of the students are Biology majors, 43% of the students are sophomores, and 32% of the sophomores are Biology majors. If a student in this class is selected at random, approximately what is the probability that the selected student is a sophomore majoring in Biology?

(A) 0.08

(B) 0.14

(C) 0.32

(D) 0.37

(E) 0.74

35. In a large city, researchers selected a random sample of 120 individuals who have multiple televisions in their house connected to a streaming device. Each individual was asked how many hours of streaming content they watch in a typical week. The sample had a mean of 12.5 hours. For which of the following populations is 12.5 hours a reasonable estimate of the mean hours of weekly streaming content watched?

(A) All individuals who regularly watch streaming content

(B) All individuals from this city who regularly watch streaming content

(C) All individuals from this city with multiple streaming devices in their house

(D) All individuals from this city with multiple streaming devices in their house who regularly watch streaming content

(E) All 120 individuals in this city with multiple streaming devices in their home who regularly watch streaming content

---

36. Suppose that at a large university, 40% of the males and 50% of the females voted in the last U.S. presidential election. José plans to take a random sample of 100 of the males and compute the sample proportion who voted. Independently, Anya plans to take a random sample of 150 of the females and compute the sample proportion who voted. What is the standard deviation of the distribution of all possible differences in the sample proportions José and Anya could obtain?

(A) $\frac{45}{\sqrt{250}}$

(B) $\sqrt{\frac{0.45(0.55)}{250}}$

(C) $\sqrt{\frac{40^2}{100}+\frac{50^2}{150}}$

(D) $\sqrt{\frac{0.40(0.60)}{100}+\frac{0.50(0.50)}{150}}$

(E) $\sqrt{\frac{0.40(0.60)}{100}-\frac{0.50(0.50)}{150}}$

37. The results from a regression analysis of data collected on SAT score and current GPA for a random sample of 610 students at a large university are summarized in the partial computer output below.

Term	Coef	SE Coef
Constant	1.406	0.143
SAT	0.001340	0.000114

$r = 0.431$ $R^2 = 0.1862$

Assume all conditions for inference were met. Based on the output shown, is there a statistically significant linear relationship between SAT score and current GPA for all students at this university?

(A) No, because the estimated slope is very close to 0.

(B) No, because the correlation is not close to -1 or 1.

(C) No, because R^2 is quite small.

(D) Yes, because the ratio of 1.406 to 0.143 is very large.

(E) Yes, because the ratio of 0.001340 to 0.000114 is very large.

38. A large insurance company sells homeowner's insurance to compensate homeowners in case of damage to their home. Suppose the insurance company's annual payout on a single policy has a severely right-skewed distribution with a mean of $4750 and a standard deviation of $2150. Which of the following best states the Central Limit Theorem in this context?

(A) The distribution of annual payouts for all such policies will be approximately normal if the population size is sufficiently large.

(B) The distribution of annual payouts in a random sample of these policies will be approximately normal if the sample size is sufficiently large.

(C) The distribution of the mean annual payout for many random samples of policies of the same size will be approximately normal if the sample size is sufficiently large.

(D) The value of the mean annual payout for a random sample of these policies will be equal to $4,750 if the sample size is sufficiently large.

(E) The variability of annual payouts in a random sample of these policies will be smaller for larger sample sizes.

---

39. The campaign manager for a presidential candidate believes that there is a significant difference in the favorability of her candidate among female and male voters. She randomly selected 300 likely female voters and 250 likely male voters and found that 144 of the females and 124 of the males favor her candidate. Given the 95% confidence interval of (-0.10, 0.07), which of the following conclusions is most appropriate?

(A) With 95% confidence, the data suggest that there is a difference between the proportions of females and males that favor the candidate since the observed difference in the proportion who favor the candidate is negative.

(B) With 95% confidence, the data suggest that there is a difference between the proportions of females and males that favor the candidate since 0 is included in the interval.

(C) With 95% confidence, the data do not suggest that there is a difference between the proportions of females and males that favor the candidate since the observed difference in the proportion who favor the candidate is negative.

(D) With 95% confidence, the data do not suggest that there is a difference between the proportions of females and males that favor the candidate since 0 is not included in the interval.

(E) With 95% confidence, the data do not sugest that there is a difference between the proportions of females and males that favor the candidate since 0 is included in the interval.

40. The Department of Transportation (DOT) is responsible for replacing road signs that are no longer adequately reflective. The DOT has replacement signs in a warehouse but wants to know whether their inventory differs from the types of signs found in a certain area. The DOT takes a random sample of 100 signs and observes the type of each. The sample data are tabulated below along with the known percentage of each type currently in inventory.

	Sign Type			
	Route	Stop	Yield	Other
Sample Count	37	45	9	9
Inventory %	36%	44%	9%	11%

Of the following, which is the best procedure to investigate whether there is convincing statistical evidence that the distribution of sign types in this area differs from the distribution in the DOT's inventory?

(A) A chi-square goodness-of-fit test

(B) A chi-square test of independence

(C) A chi-square test of homogeneity

(D) A matched-pairs *t*-test for a mean difference

(E) A two-sample *t*-test for the difference between two means

END OF EXAMINATION