Table of z values and probabilities for the standard normal distribution. z is the first column plus the top row. Each cell shows $P(X \leq z)$. For example $P(X \leq 1.04) = .8508$. For $z < 0$ subtract the value from 1, e.g., $P(X \leq -1.04) = 1 - .8508 = .1492$.

z	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990

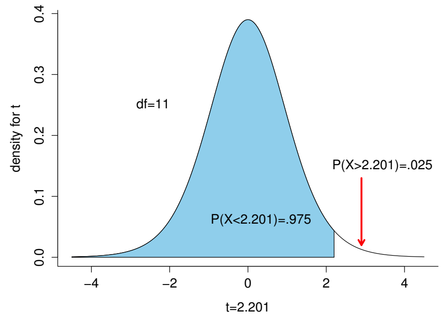


Table of t values and right tail probabilities. Degrees of freedom are in the first column (df). **Right tail probabilities** are in the first row. For example for $d.f. = 7$ and $\alpha = .05$ the critical t value for a two-tail test is 2.365 and for $d.f. = 10$ and $\alpha = .1$ the critical t value for a one-tail test is 1.372.

df	.1	.05	.025	.01	.005
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
30	1.310	1.697	2.042	2.457	2.750
40	1.303	1.684	2.021	2.423	2.704
50	1.299	1.676	2.009	2.403	2.678
75	1.293	1.665	1.992	2.377	2.643
100	1.290	1.660	1.984	2.364	2.626

Table of F values for right tail probabilities of .05. Degrees of freedom for denominator are in the first column and degrees of freedom for the numerator are in the top row.

denom.	numerator df									
df	1	2	3	4	5	7	10	15	50	100
1	161.45	199.5	215.71	224.58	230.16	236.77	241.88	245.95	251.77	253.04
2	18.51	19	19.16	19.25	19.3	19.35	19.4	19.43	19.48	19.49
3	10.13	9.55	9.28	9.12	9.01	8.89	8.79	8.7	8.58	8.55
4	7.71	6.94	6.59	6.39	6.26	6.09	5.96	5.86	5.7	5.66
5	6.61	5.79	5.41	5.19	5.05	4.88	4.74	4.62	4.44	4.41
6	5.99	5.14	4.76	4.53	4.39	4.21	4.06	3.94	3.75	3.71
7	5.59	4.74	4.35	4.12	3.97	3.79	3.64	3.51	3.32	3.27
8	5.32	4.46	4.07	3.84	3.69	3.5	3.35	3.22	3.02	2.97
9	5.12	4.26	3.86	3.63	3.48	3.29	3.14	3.01	2.8	2.76
10	4.96	4.1	3.71	3.48	3.33	3.14	2.98	2.85	2.64	2.59
11	4.84	3.98	3.59	3.36	3.2	3.01	2.85	2.72	2.51	2.46
12	4.75	3.89	3.49	3.26	3.11	2.91	2.75	2.62	2.4	2.35
13	4.67	3.81	3.41	3.18	3.03	2.83	2.67	2.53	2.31	2.26
14	4.6	3.74	3.34	3.11	2.96	2.76	2.6	2.46	2.24	2.19
15	4.54	3.68	3.29	3.06	2.9	2.71	2.54	2.4	2.18	2.12
16	4.49	3.63	3.24	3.01	2.85	2.66	2.49	2.35	2.12	2.07
17	4.45	3.59	3.2	2.96	2.81	2.61	2.45	2.31	2.08	2.02
18	4.41	3.55	3.16	2.93	2.77	2.58	2.41	2.27	2.04	1.98
19	4.38	3.52	3.13	2.9	2.74	2.54	2.38	2.23	2	1.94
20	4.35	3.49	3.1	2.87	2.71	2.51	2.35	2.2	1.97	1.91
21	4.32	3.47	3.07	2.84	2.68	2.49	2.32	2.18	1.94	1.88
22	4.3	3.44	3.05	2.82	2.66	2.46	2.3	2.15	1.91	1.85
23	4.28	3.42	3.03	2.8	2.64	2.44	2.27	2.13	1.88	1.82
24	4.26	3.4	3.01	2.78	2.62	2.42	2.25	2.11	1.86	1.8
25	4.24	3.39	2.99	2.76	2.6	2.4	2.24	2.09	1.84	1.78
26	4.23	3.37	2.98	2.74	2.59	2.39	2.22	2.07	1.82	1.76
27	4.21	3.35	2.96	2.73	2.57	2.37	2.2	2.06	1.81	1.74
28	4.2	3.34	2.95	2.71	2.56	2.36	2.19	2.04	1.79	1.73
29	4.18	3.33	2.93	2.7	2.55	2.35	2.18	2.03	1.77	1.71
30	4.17	3.32	2.92	2.69	2.53	2.33	2.16	2.01	1.76	1.7
40	4.08	3.23	2.84	2.61	2.45	2.25	2.08	1.92	1.66	1.59
60	4	3.15	2.76	2.53	2.37	2.17	1.99	1.84	1.56	1.48
100	3.94	3.09	2.7	2.46	2.31	2.1	1.93	1.77	1.48	1.39
1000	3.85	3	2.61	2.38	2.22	2.02	1.84	1.68	1.36	1.26

Chapter 3:

sample mean: $\bar{x} = \sum_{i=1}^{n} x_i/n$

sample variance: $s^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}.$

sample standard deviation: $s = \sqrt{s^2}.$

Coefficient of Variation: $CV = \frac{s}{\bar{x}}\left(100\%\right)$

sample z-Score: $z = \frac{x_i - \bar{x}}{s}$

Interquartile Range: IQR $= Q_3 - Q_1$.

Sample Covariance: $s_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{n-1}$

Sample Correlation Coefficient: $r_{xy} = s_{xy}/(s_x s_y)$

Chapter 4:

The complement rule: $P(A) + P(A') = 1$

addition rule: $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$

conditional probability: $P(A|B) = \frac{P(A \text{ and } B)}{P(B)}$

Bayes' Theorem

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{P(A_1)P(B|A_1) + P(A_2)P(B|A_2) + \cdots + P(A_n)P(B|A_n)}$$

Combinations: ${}_nC_x = \frac{n!}{(n-x)!x!}$

Chapter 5:

Expected Value and mean of a Discrete Probability Distribution:

$$E(x) = \mu = \sum_{i=1}^{n} x_i P(x_i)$$

Variance of a Discrete Probability Distribution:

$$\sigma^2 = \sum_{i=1}^{n}(x_i - \mu)^2 P(x_i)$$

Chapter 5 continued:

Binomial Probability Dist.: $P(x,n) = \frac{n!}{(n-x)!x!}p^x(q)^{(n-x)}$

Mean of a Binomial Distribution: $\mu = np$

Standard Dev. of a Binomial Distribution: $\sigma = \sqrt{npq}$

Poisson Probability Distribution: $P(x) = \frac{\lambda^x e^{-\lambda}}{x!}$

Chapter 6:

Normal PDF: $f(x) = \frac{1}{\sigma\sqrt{2\pi}}e^{-(1/2)[(x-\mu)/\sigma]^2}$

the z-score: $z = \frac{x-\mu}{\sigma}$

Exponential PDF: $f(x) = \lambda e^{-\lambda x}$

Exponential CDF: $P(x \le a) = 1 - e^{-a\lambda}$

Standard Dev. of Exponential Dist.: $\sigma = \mu = \frac{1}{\lambda}$

Continuous Uniform PDF

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{if } a \le x \le b \\ 0 & \text{otherwise} \end{cases}$$

Uniform CDF: $P(x_1 \le x \le x_2) = \frac{x_2 - x_1}{b-a}$

mean of the continuous uniform dist.: $\mu = \frac{a+b}{2}$

standard dev. of the continuous uniform dist.: $\sigma = \frac{b-a}{\sqrt{12}}$

Chapter 7:

standard error of the mean: $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}.$

z-score for the mean: $z_{\bar{x}} = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}}$

Chapter 7 continued:

sample proportion: $\bar{p} = \frac{x}{n}$

standard error of the proportion: $\sigma_{\bar{p}} = \sqrt{\frac{p(1-p)}{n}}$

Chapter 8:

Confidence Interval for the mean (σ known):

$$\bar{x} \pm z_{\alpha/2}\sigma_{\bar{x}}$$

margin of error for a CI for the mean: $ME_{\bar{x}} = z_{\alpha/2}\sigma_{\bar{x}}$

approximate standard error of the mean: $\hat{\sigma}_{\bar{x}} = \frac{s}{\sqrt{n}}.$

Confidence Interval for the mean (σ unknown):

$$\bar{x} \pm t_{\alpha/2}\hat{\sigma}_{\bar{x}}$$

Sample Size needed to Estimate a population mean

$$n = \frac{(z_{\alpha/2})^2\sigma^2}{(ME_{\bar{x}})^2}$$

Sample Size needed to Estimate the population proportion

$$n = \frac{(z_{\alpha/2})^2\bar{p}(1-\bar{p})}{(ME_p)^2}$$

Chapter 9:

the z-test statistic for a hypothesis test for the population mean (when σ is known)

$$z_{\bar{x}} = \frac{\bar{x} - \mu_{H_0}}{\sigma/\sqrt{n}}$$

the t-test statistic for a hypothesis test for the population mean (when σ is unknown)

$$t_{\bar{x}} = \frac{\bar{x} - \mu_{H_0}}{s/\sqrt{n}}$$

Chapter 10:
the mean of the sampling distribution for the difference in means:

$$\mu_{\bar{x}_1-\bar{x}_2} = \mu_{\bar{x}_1} - \mu_{\bar{x}_2}$$

the standard error of the difference between two means:

$$\sigma_{\bar{x}_1-\bar{x}_2} = \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

the z-test statistic for a hypothesis test for the difference between two means (σ_1 and σ_2 known)

$$z_{\bar{x}} = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)_{H_0}}{\sigma_{\bar{x}_1-\bar{x}_2}}$$

the t-test statistic for a hypothesis test for the difference between two means (σ_1 and σ_2 unknown but equal)

$$t_{\bar{x}} = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)_{H_0}}{\sqrt{s_p^2\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

pooled variance: $s_p^2 = \dfrac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{(n_1-1)+(n_2-1)}$

the t-test statistic for a hypothesis test for the difference between two means (σ_1 and σ_2 unknown and unequal)

$$t_{\bar{x}} = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)_{H_0}}{\sqrt{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)}}$$

Confidence Interval for the difference between the means of two independent populations (σ_1 and σ_2 unknown but equal)

$$(\bar{x}_1 - \bar{x}_2) \pm t_{\alpha/2}\sqrt{\frac{s_p^2}{n_1} + \frac{s_p^2}{n_2}}$$

the matched-pair difference: $d = x_1 - x_2$

the mean of matched-pair difference: $\bar{d} = \dfrac{\sum_{i=1}^n d_i}{n}$

the standard deviation of the matched-pair differences

$$s_d = \sqrt{\frac{\sum_{i=1}^n (d_i - \bar{d})^2}{n-1}}$$

the t-Test Statistic for a Matched-Pair hypothesis test for the mean

$$t_{\bar{x}} = \frac{\bar{d} - (\mu_d)_{H_0}}{s_d/\sqrt{n}}$$

Chapter 11:

the total sum of squares (SST): $SST = \sum_{j=1}^{k}\sum_{i=1}^{n_j}(x_{ij} - \bar{\bar{x}})^2$

the mean square total (MST): $MST = \dfrac{SST}{n_T - 1}$

the partitioning of the Total Sum of Squares (SST) for a One-Way ANOVA: SST=SSB+SSW.

sum of squares between (SSB): $SSB = \sum_{j=1}^{k} n-j(\bar{x}_j - \bar{\bar{x}})^2$

the mean square between (MSB): $MSB = \dfrac{SSB}{k-1}$

sum of squares within (SSW): $SSW = \sum_{j=1}^{k}\sum_{i=1}^{n_j}(x_{ij} - \bar{x}_j)^2$

the mean square within (MSW): $MSW = \dfrac{SSW}{n_T - k}$

the F-test statistic for One-Way ANOVA: $F_{\bar{x}} = \dfrac{MSB}{MSW}$

Tukey-Kramer critical range:

$$CR_{ij} = Q_\alpha\sqrt{\frac{MSW}{2}\left(\frac{1}{n_i} + \frac{1}{n_j}\right)}$$

Chapter 14: simple linear regression model for a population $y_i = \beta_0 + \beta_1 x_i + \epsilon_i$

$$\hat{y} = b_0 + b_1 x \qquad e_i = y_i - \hat{y}_i$$

sum of squares error: $SSE = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2$

total sum of squares (SST): $SST = \sum(y - \bar{y})^2$

$$SST = SSR + SSE$$

sum of squares regression (SSR): $SSR = \sum(\hat{y} - \bar{y})^2$

Chapter 14 continued:

$$R^2 = \frac{SSR}{SST}$$

F-statistic for the coef. of determination: $F = \dfrac{SSR}{SSE/(n-2)}$

Standard Error of the Estimate, $s_e = \sqrt{SSE/(n-2)}$.
Confidence Interval (CI) for an average value of Y:

$$CI = \hat{y}^* \pm t_{\alpha/2}s_e\sqrt{\frac{1}{n} + \frac{(x-\bar{x})^2}{\sum x^2 - ((\sum x)^2/n)}}$$

Prediction Interval (PI) fpr a specific value of y:

$$PI = \hat{y}^* \pm t_{\alpha/2}s_e\sqrt{1 + \frac{1}{n} + \frac{(x-\bar{x})^2}{\sum x^2 - ((\sum x)^2/n)}}$$

t-test statistic for the regression slope: $t = \dfrac{b_1 - \beta_1}{s_b}$

the standard error of a slope: $s_b = \dfrac{s_e}{\sqrt{\sum x^2 - n(\bar{x})^2}}$

confid. interval for the pop. slope: $CI = b_1 \pm t_{\alpha/2}s_b$

Chapter 15:
mean square regression (MSR): $MSR = SSR/k$

mean square error (MSE): $MSE = SSE/(n-k-1)$

F-test stat. for the overall regression model: $F = \dfrac{MSR}{MSE}$

adjusted multiple coef. of det.: $R_A^2 = 1 - (1-R^2)\dfrac{n-1}{n-k-1}$

variance inflation factor: $VIF_j = \dfrac{1}{1-R_j^2}$

Other Math Rule Reminders:

$$e^x = \exp(x) \quad \text{and} \quad \ln 1 = 0 \quad \text{and} \quad \ln e = 1$$

$$x! = (x)(x-1)(x-2)\cdots(2)(1) \quad \text{and} \quad 0! = 1 \quad \text{and} \quad x^0 = 1$$

ECN221 Exam 2 B FALL 2016 (Chapters 1-8 Donnelly), ASU-COX

NAME:____

VERSION B

Choose the best answer. Do not write letters in the margin or communicate with other students in any way. If you have a question note it on your exam and ask for clarification when your exam is returned. In the meantime choose the best answer. Neither the proctors nor Dr. Cox will answer questions during the exam.

Please check each question and possible answers thoroughly as questions at the bottom of a page sometimes run onto the next page. Please verify that your test version and scantron version are the same.

This exam has 25 questions. Cost for turning in exam late: 1st minute is 10 points; 2nd minute is 20 additional points (30 total); 3rd minute is 30 points (60 total) ; 4th minute is 40 points (100 total); 5th minute is 50 points (150 total); 6th minute is 80 points (230 total) and no exams are graded past that point.

1. I have checked that my ID is bubbled in correctly. If it is bubbled in incorrectly I will get this question wrong. I also understand that questions and their possible answers may run onto the next page and so I should always check the top of the next page for possible answers. I understand that if I have a question I should simply make a note on my exam and ask Dr. Cox afterwards. I should always choose the best answer.

   (a) False.
   (b) I didn't read the directions.
   (c) True.

2. The Chainsmokers gave a recent show in Las Vegas at the Hakkasan night club (Oct 7, 2016). The ticket prices for VIP packages with prepurchased drinks were (net of alcohol) \$25 for women and \$125 for men. The club has a capacity of around 7,000 (though not all of that is in the main room where The Chainsmokers djed). Suppose that 5,000 tickets were sold for the event and 60% of those were to women. What was the average ticket price in dollars?

   (a) 61
   (b) 71.25
   (c) 65
   (d) 63.75

3. What is the probability of observing a z value less than -0.612?

   (a) 0.7297
   (b) 0.3903
   (c) 0.2003
   (d) 0.2703

4. α is known as the confidence level.

   (a) True
   (b) False

5. We can define outliers as observations at least three standard deviation away from the mean, i.e. $|z| \geq 3$. We can also define outliers as observations that are more than 1.5 IQR below Q1 or above Q3. True or false: In any data set there are always at least as many outliers using the z-score definition as there are using the IQR definition.

   (a) True.

(b) False.

(c) True if all the data values are positive but not necessarily true if negative values are possible.

6. True or false: No data set can have more than one observation with a z-value greater than 10.

   (a) True.

   (b) False.

7. The graph below (possibly on the next page), for the "Presidential Poll" shows clearly that

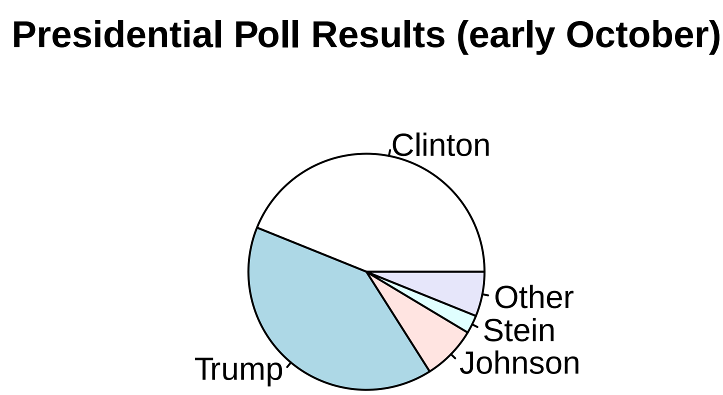
   (a) Clinton is polling at over 50%.

   (b) Johnson is polling higher than Stein.

   (c) Justice Ginsburg will leave the country.

   (d) Bar charts can be difficult to read.

   (e) Clinton is polling higher than Trump.



8. Using the Poisson PMF, when the mean is 12, what is $P(X = 10)$?

(a) 0.3072559
(b) 0.0329024
(c) 0.8951627
(d) 0.1048373

9. Suppose that the average number of online purchases a customer makes in a week is 0 and that for each customer this random variable follows the Poisson distribution. What is the probability that a customer will make exactly 1 online purchases in a week?

   (a) 0
   (b) 0.1
   (c) -0.03
   (d) 1

10. Suppose at ASU 46% of students take out some form of student loan to pay for school. In a class of 383 what is the expected number of students with a loan?

   (a) 186.18
   (b) 146.18
   (c) 176.18
   (d) 153.2766

11. You are tasked with finding out about ASU student attitudes about the upcomming election. You have access to student records and contact information. You decide to email a survey to a selected group of students and offer them a gift card for participating in the survey. To choose to whom you will send the email you have a computer program sort the students alphabetically by last name and select every 50th name. This is an example of

   (a) convenience sampling
   (b) random sampling
   (c) stratified sampling
   (d) cluster sampling
   (e) systematic sampling

12. Suppose you have a random variable that is uniformly distributed with a maximum of 263 and a minimum of 83. What is the expected value of this random variable?

(a) 173
(b) 193
(c) 158
(d) 224.9

13. Suppose that starting salaries for new graduates are roughly exponentially distributed. That is, salaries tend to group around a low end and then are skewed right with a few "rock star" students earning high salaries. If the average salary is \$56651 then what is the probability of getting a salary of at least \$65000?

(a) 0.3174687
(b) 0.6825313
(c) 0.1587344
(d) 0.6349374

14. What is the critical t value when there are 15 degrees of freedom and the confidence coefficient is .98?

(a) 2.947
(b) 2.602
(c) 1.341
(d) 1.753
(e) 2.131

15. Suppose that the known standard deviation for the numbers of hours that students work in a week is 15. If I draw a sample of 42 what is the standard error?

(a) 2.3145502
(b) 1.8516402
(c) 0.3571429
(d) 3.0089153

16. A confidence interval is constructed as $\bar{x}\pm$ margin of error.

(a) true
(b) false
(c) true if $\alpha < .5$ and false if $\alpha \geq .5$

17. Suppose that in a recent sample of 29 recent graduates with business degrees the mean starting salary is 52287 and the standard deviation is 1512. Construct a 99% confidence interval for the mean starting salary for business majors.

    (a) [51511, 53062]
    (b) [51911, 52662]
    (c) [50911, 53662]
    (d) [46359, 47755]

18. You are given the following, $s = 8$, $\bar{x} = 56$ and $n = 23$. Construct a 90% confidence interval. The resulting interval is

    (a) [53.1356052, 58.8643948]
    (b) [51.1356052, 62.8643948]
    (c) [47.1356052, 64.8643948]
    (d) [47.8220447, 61.8076146]

19. Suppose that you have a sample with 27 observations. You are going to use this sample to construct a confidence interval for the population mean. How many degrees of freedom are there?

    (a) 5.1961524
    (b) 27
    (c) 14
    (d) 26

20. For a confidence coefficient of .7 we construct confidence intervals for 50 samples. Approximately how many of our confidence intervals will contain the true parameter?

    (a) 35
    (b) 30
    (c) 25
    (d) 15

21. Which of the following is correct. Let the confidence coefficient be .95 and suppose we are going to draw a random sample and construct the appropriate confidence interval.

(a) There is a 95% chance that the true parameter will fall within our confidence interval.

(b) There is a 5% chance that the true parameter will fall within our confidence interval.

(c) There is a 95% chance that we draw a sample for which the corresponding confidence interval will contain the true parameter.

(d) There is a 5% chance that we draw a sample for which the corresponding confidence interval will contain the true parameter.

22. One of the assigned readings discussed some of the contributions of Abraham Wald to statistics. We might think of one of these contributions as being

(a) demonstrating the effect of varying the sample size on measures of central tendency.

(b) discovering the binomial distribution.

(c) developing a formula for the calculation of the standard deviation.

(d) demonstrating the importance of considering units that don't end up in our data sets.

(e) demonstrating the importance of having a sufficiently large sample.

23. Here are the top pitchers by WAR for 2016 from fangraphs (www.fangraphs.com). What is the standard deviation of wins?

2016 Fangraphs Top Pitcher WAR

Name	wins	innings	strikeouts
Kershaw	12	149	172
Syndergaard	14	183.67	218
Fernandez	16	182.33	253
Scherzer	20	228.33	284
Cueto	18	219.67	198
Porcello	22	223	189
Verlander	16	227.67	254
Sale	17	226.67	233

(a) 1.09

(b) 3.51

(c) 2.77

(d) 6.37

(e) 3.18

24. Suppose that the data below are the *entire population*. **Do not** recalculate the standard deviation; suppose that the value you found in the previous problem is the population standard deviation. Suppose we draw a sample of Kershaw and Cueto. What is the z-value for this sample? Note: the variable of interest is the number of wins.

2016 Fangraphs Top Pitcher WAR

Name	wins	innings	strikeouts
Kershaw	12	149	172
Syndergaard	14	183.67	218
Fernandez	16	182.33	253
Scherzer	20	228.33	284
Cueto	18	219.67	198
Porcello	22	223	189
Verlander	16	227.67	254
Sale	17	226.67	233

(a) .233
(b) -1.033
(c) -.633
(d) -.833

25. Again you will use the data pitchers for 2016. This time look at the variable "strike-outs." Which pitcher has the largest z score in absolute terms (could be positive or negative)?

2016 Fangraphs Top Pitcher WAR

Name	wins	innings	strikeouts
Kershaw	12	149	172
Syndergaard	14	183.67	218
Fernandez	16	182.33	253
Scherzer	20	228.33	284
Cueto	18	219.67	198
Porcello	22	223	189
Verlander	16	227.67	254
Sale	17	226.67	233

(a) Scherzer

(b) Syndergaard
(c) Cueto
(d) Verlander
(e) Sale

Key

1. c
2. c, .6*25+.4*125 or you could have added up 25 3,000 times and 125 2,000 times and then divided by 5,000 but that would take a lot longer.
3. d
4. b, go back and check the lecture notes
5. b, consider an example where the observations are -3, -2, -2, -2, -1.9, -1.9, -1.9, -1.9, 0.
6. b, consider a data set with 999,000 zeros and 2 observations of 1,000,000. The mean is essentially 2 and the standard deviation is around 2,828.43. But more simply, when there are a large number of observations, as in hundereds of thousands or millions, it is reasonable to observe multiple values with large z-scores.
7. b
8. d, plug in to the Poisson PMF which is on the formula sheet
9. a
10. c
11. e, check definitions in the text and lecture notes
12. a
13. a, the formula for the exponential distribution is on the formula sheet
14. b
15. a, the standard error is $\sigma/\sqrt{n}$
16. a, construct the interval as $\bar{x}\pm$ margin of error.
17. a
18. a, when using the sample standard deviation find a t value instead of a z value.
19. d
20. a, follows from the definition of confidence coefficient
21. c, remember that the true parameter is fixed.

22. d
23. e
24. d, find the average for those two pitchers and subtract the overall average and divide by the standard error $3.18/\sqrt{2}$.
25. a