Solutions for Practice Exam of Statistics

1. **You're manufacturing a widget and using an X-bar and R chart to control the critical feature of the product. Your normal process has the following attributes:** **X-double bar is 225, R-bar is 12, n = 8.**
   Identify the lower control limits for the X-bar chart:
   - **220.52**
   - 229.48
   - 233.14
   - 218.71

X-Bar and R Chart				
Subgroup Sample Size	X-Bar Factor	Range Factors		Variance Factor
n	A_2	D_3	D_4	d_2
2	1.880	-	3.267	1.128
3	1.023	-	2.575	1.693
4	0.729	-	2.282	2.059
5	0.577	-	2.115	2.326
6	0.483	-	2.004	2.534
7	0.419	0.076	1.924	2.704
8	0.373	0.136	1.864	2.847
9	0.337	0.184	1.816	2.970
10	0.308	0.223	1.777	3.078
15	0.223	0.347	1.653	3.472
20	0.180	0.415	1.585	3.735
25	0.153	0.459	1.541	3.931

$$\boldsymbol{Lower\ Control\ Limit}: \boldsymbol{LCL_{\overline{X}}} = \bar{\bar{X}} - A_2\bar{R}$$

Now we must look up the A_2 constant using the sample size (n=8), and we find A_2 = 0.373.

$$\boldsymbol{LCL_{\overline{X}}} = 225 - 0.373 * 12 = 220.52$$

2. **Calculate Cpk for the following Parameters: (USL = 15, LSL = 10, μ = 13, σ = 1.25)**
   - **0.53**
   - 0.67
   - 0.80
   - 1.0

$$C_{pk} = Min(C_{p,Lower}, C_{p,Upper}) = Min\left(\frac{USL - \tilde{x}}{3s}, \frac{\tilde{x} - LSL}{3s}\right)$$

$$C_{pk} = Min\left(\frac{15-13}{3*1.25}, \frac{13-10}{3*1.25}\right) = Min\left(\frac{2}{3.75}, \frac{3}{3.75}\right) = Min(0.53\ , 0.80) = \mathbf{0.53}$$

3. **You work at a shipping facility whose shipping failures follow the Poisson distribution. You ship approximately 1,000 packages per day, and the mean number of shipping errors is equal to 15 per day. What is the probability that you will experience exactly 15 failures in one day?**
   - 5%
   - **10%**
   - 15%
   - 25%

$$f(x) = P(X = x) = \frac{e^{-\lambda} * \lambda^x}{x!}$$

$$f(15) = P(X = 15) = \frac{e^{-15} * 15^{15}}{15!} = 0.1024 = 10.2\%$$

4. **What is the critical z-value associated with a 2-sided confidence interval that's associated with a 1% alpha risk?**
   - **z-score = 2.58**
   - z-score = 2.33
   - z-score = 1.96
   - z-score = 3.09

   Because it's a 2-sided distribution, we're looking for the z-score that's associated with the area under the curve of 0.495. This would capture 49.5% on the left half & right half of the distribution, leaving the remaining 1% of the alpha risk in the rejection area of the tails of the distribution.

   The z-score associated with 0.495 probability is z = 2.58

Area under the Normal Curve from 0 to X										
X	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.00000	0.00399	0.00798	0.01197	0.01595	0.01994	0.02392	0.02790	0.03188	0.03586
0.1	0.03983	0.04380	0.04776	0.05172	0.05567	0.05962	0.06356	0.06749	0.07142	0.07535
2.2	0.48610	0.48645	0.48679	0.48713	0.48745	0.48778	0.48809	0.48840	0.48870	0.48899
2.3	0.48928	0.48956	0.48983	0.49010	0.49036	0.49061	0.49086	0.49111	0.49134	0.49158
2.4	0.49180	0.49202	0.49224	0.49245	0.49266	0.49286	0.49305	0.49324	0.49343	0.49361
2.5	0.49379	0.49396	0.49413	0.49430	0.49446	0.49461	0.49477	0.49492	0.49506	0.49520
2.6	0.49534	0.49547	0.49560	0.49573	0.49585	0.49598	0.49609	0.49621	0.49632	0.49643
2.7	0.49653	0.49664	0.49674	0.49683	0.49693	0.49702	0.49711	0.49720	0.49728	0.49736

5. **For many of the tools within Inferential Statistics there are assumptions that must be made. What assumption is described as the assumption that all of the sample groups being analyzed have the same variance between the groups.**
   - The Assumption of Random Sampling
   - **The Assumption of Homogeneous Variances**
   - The Assumption of Linearity
   - The Assumption of Normality

6. **You manufacture a widget and use a c chart to monitor the number of defects associated with your process. Your sample size is constant and on average you find 7 defects per sample. Identify the upper control limits for the c chart:**
   - 4
   - 10
   - 12
   - **15**

$$UCL_c = \bar{c} + 3\sqrt{\bar{c}} = 7 + 3\sqrt{7} = \mathbf{14.94}$$

7. **You're performing a hypothesis test for the population mean, and your sample mean is 2.53, your null hypothesis for the population mean is 2.50, your sample size is 50 and your population standard deviation is 0.10. Calculate your z test statistic:**

- 0.300
- 1.732
- **2.121**
- 2.460

In this instance our hypothesis test sample size is greater than 30 and we know the population standard deviation; therefore we can use the normal distribution and z-score for our test statistic.

$$z = \frac{\bar{x} - \mu}{\frac{\sigma}{\sqrt{n}}} = \frac{2.53 - 2.50}{\frac{0.10}{\sqrt{50}}} = 2.121$$

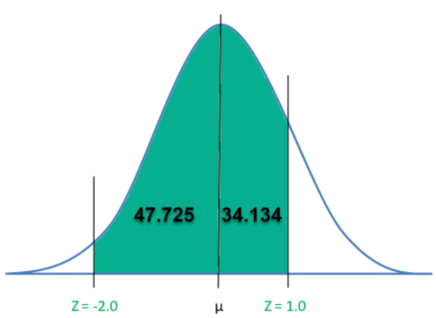
8. **For a random variable that is normally distributed with a mean value of 10 and a variance of 4, what is the probability of occurrence of the values between 6 and 12.**

- 47%
- 75%
- 66%
- **82%**

To solve this problem, we first must calculate the Z-score for both 1.25 & 2.00.

$$Z = \frac{6 - 10}{2} = -2.0 \qquad AND \qquad Z = \frac{12 - 10}{2} = 1.0$$

Then we need to go to the table to look up the probability of Z = 1.0 & Z = 2.0 and add them up. the Probability of Z = 1.00 is 34.134% (0.34134) and the Probability of Z = 2.00 is equal to 47.725% (0.47725).



Area under the Normal Curve from 0 to X

X	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07
0.0	0.00000	0.00399	0.00798	0.01197	0.01595	0.01994	0.02392	0.02790
0.9	0.31594	0.31859	0.32121	0.32381	0.32639	0.32894	0.33147	0.33398
1.0	0.34134	0.34375	0.34614	0.34849	0.35083	0.35314	0.35543	0.35769
1.1	0.36433	0.36650	0.36864	0.37076	0.37286	0.37493	0.37698	0.37900
1.9	0.47128	0.47193	0.47257	0.47320	0.47381	0.47441	0.47500	0.47558
2.0	0.47725	0.47778	0.47831	0.47882	0.47932	0.47982	0.48030	0.48077

Graphically this looks like this, with 34.134% of the distribution existing on the left half of the shaded area, and the other 47.725% existing on the left half of the shaded area out to Z = -2.0.

When we add these two shaded areas together, we find that the area under the curve equals **81.859%.**

9. **You're preparing for an upcoming production run where the likelihood (Probability) of defect A is known to be 3%, and the likelihood of defect B is 3%; and an overlapping 1% had both defect A & defect B. If you randomly sampled 1 piece from a lot of 100, what is the likelihood of picking a defect?**
   - 4%
   - **5%**
   - 6%
   - 7%

   In this situation the following information is true: **P(A) = 3%** **P(B) = 3%** **P(A ∩ B) = 1%**
   Therefore, the probability of picking one of the two defect types can be calculated using the addition rule:

   P(A U B) = P(A) + P(B) - P(A ∩ B) = 3% + 3% - 1% = 5%

10. **How many treatments would be required for a DOE with 8 factors where a quarter factorial design is chosen:**
    - 256
    - 128
    - **64**
    - 32

$$Quarter\ Factorial\ Design: Number\ of\ Treatments = \frac{Levels^{Factors}}{4} = \frac{L^F}{4} = \frac{2^F}{2^2} = 2^{8-2} = 2^6 = 64$$

11. **You're creating a linear regression model for your data and you've calculated the following values:**

    S_{yy} = 102, S_{xy} = 168, S_{xx} = 142

    What is the slope coefficient for your regression model?
    - 0.61
    - 0.85
    - **1.18**
    - 1.39

    The slope of the linear regression model, $\beta_1 = S_{xy} / S_{xx} = 168 / 142 = 1.18$

12. **Which Statement below regarding the Central Limit Theorem is true?**
    - The Central Limit Theorem cannot be used if the distribution being sampled from is not normally distributed.
    - The Central Limit Theorem is only used in descriptive statistics, not inferential statistics
    - **The Central Limit Theorem can be applied even if you're sampling from a distribution that is not normally distributed.**
    - The Central limit theorem does not apply to hypothesis testing

13. Calculate C_{pk} for the following Parameters: (USL = 1.005, LSL = 0.950, σ = 0.010, μ =0.970)

- 0.50
- **0.67**
- 1.0
- 1.33

$$C_{pk} = Min(C_{p,Lower}, C_{p,Upper}) = Min\left(\frac{USL - \tilde{x}}{3s}, \frac{\tilde{x} - LSL}{3s}\right)$$

$$C_{pk} = Min\left(\frac{1.005 - 0.970}{3 * 0.010}, \frac{0.097 - 0.095}{3 * 0.010}\right) = Min\left(\frac{.035}{.030}, \frac{.020}{0.030}\right)$$

$$C_{pk} = Min(1.16, 0.67) = \mathbf{0.67}$$

14. You manufacture a widget and use an x-bar and S chart to monitor your process, where you sample 5 units in each subgroup, and s-bar = 4.2. Estimate the population standard deviation for this process.

- 4.2
- 2.1
- 3.9
- **4.5**

We divide S-bar by the factor c_4, which is based on the n=5 sample size.

$$\boldsymbol{Population\ Standard\ Deviation = \hat{\sigma} = \frac{\bar{s}}{C_4} = \frac{4.2}{0.9400} = 4.5}$$

15. Which measurement scales allow for a Median Value to be determined as a measure of Central Tendency?

A. Nominal
B. Ordinal
C. Interval
D. Ratio

- A,
- A, B
- B, C,
- **B, C, D**

16. **The one-way ANOVA Analysis below has 10 treatment groups with the total degrees of freedom of 19.**

Variation Source	Sum of Squares (SS)	Degrees of freedom (DF)	Mean Squares (MS)	F-Value
Treatment (Between)	**45**	9	5	**0.90**
Error (Within)	55	**10**	5.5	
Total	100	19		

Calculate the Treatment Mean Square for this ANOVA Table.

- 4.5
- **5**
- 5.5
- 6.1

First, we can solve for the treatment sum of squares by simply subtracting 55 from 100, to get a treatment sum of Square of **45**. Then we must solve for the degrees of freedom. The treatment degrees of freedom is equal to the number of treatment levels (10) - 1; so 9 degrees of freedom. Then we can solve for the error degrees of freedom by subtracting 19 - 9; so, **10** degrees of freedom.

Then we can calculate the mean squares for the treatment as the treatment sum of square (45) divided by the treatment degrees of freedom (9): 45/9 = 5.

17. **You're performing a hypothesis test to compare the sample variance to see if it's equivalent to a hypothesized population variance. You take 10 samples, and your hypothesis test has 10% alpha risk (2-tailed test). What is the left-tail critical value for this test?**

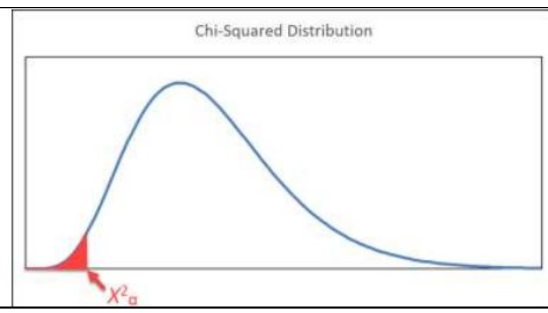
- **3.325**
- 3.940
- 4.168
- 4.865

When performing a hypothesis test to compare a sample variance against the population variance, we used the chi-squared distribution.

If we take 10 samples, our degrees of freedom:

$$d.f. = n - 1 = 9$$

With a 2-sided test, we have 5% alpha risk in the left tail, and at 9 degrees of freedom the left tail critical value is **1.735**.



Left-Tail Critical Value of the Chi-Squared (X^2) Distribution

df (v)	0.001	0.005	0.010	0.025	0.050	0.100
1	0.000	0.000	0.000	0.001	0.004	0.016
2	0.002	0.010	0.020	0.051	0.103	0.211
3	0.024	0.072	0.115	0.216	0.352	0.584
4	0.091	0.207	0.297	0.484	0.711	1.064
5	0.210	0.412	0.554	0.831	1.145	1.610
6	0.381	0.676	0.872	1.237	1.635	2.204
7	0.598	0.989	1.239	1.690	2.167	2.833
8	0.857	1.344	1.646	2.180	2.733	3.490
9	1.152	1.735	2.088	2.700	3.325	4.168
10	1.479	2.156	2.558	3.247	3.940	4.865

18. **You're creating a linear regression model for your data and you've calculated the following values:**

S_{yy} = 1125, S_{xy} = 75, S_{xx} = 5, β_0 = -12

What is the predicted value of Y when X = 10:

- 75
- **138**
- 150
- 738

$$Y(x) = \beta_0 + \beta_1 * x$$

To solve for Y, we need to calculate the slope coefficient: $\beta_1 = \frac{S_{xy}}{S_{xx}} = \frac{75}{5} = 15$

Now we can solve for Y(X = 10). $Y(x = 10) = -12 + 15 * 10 = 138$

19. **Fill in the blank: Two factors are _________________ when their effects are indistinguishably combined to affect the response variable.**

- **Confounding**
- Replicates
- Randomized
- Interacting

20. **Calculate C_r for the following Parameters: (USL = 675, LSL = 625, σ = 5)**

- **0.60**
- 0.95
- 1.00
- 1.20

$$C_r = \frac{1}{C_p} = \frac{6\sigma}{USL - LSL} = \frac{6 * 5}{675 - 625} = \frac{30}{50} = \mathbf{0.60}$$

21. **If the probability of event A is P(A) = 0.50 and the probability of event B is P(B) = .60 and the intersection of A &B is P(A & B) = 0.20, find the probability of A given that B has occurred - P(A|B).**

- 25%
- **33%**
- 40%
- 60%

P(A|B) = P(A & B) / P(B)

P(A|B) = 0.20 / 0.60 = 1/3

P(A|B) = **33%.**

22. You've sampled 60 units from the latest production lot to measure the width of the product. The sample mean is 6.75in and the population standard deviation is known to be 0.75in. Calculate the 95% confidence interval for the population mean:

- 6.75 $\pm$ 0.219
- 6.75 $\pm$ 1.470
- 6.75 $\pm$ 0.024
- **6.75 $\pm$ 0.189**

Because we've sampled more than 30 units and the population standard deviation is known, we can use the Z-score approach to this confidence interval problem. We also need to find the Z-score associated with the 95% confidence interval using the Z-Table, we find Z = 1.96.

$$Interval\ Estimate\ of\ Population\ Mean\ (known\ variance)\ :\bar{x}\ \pm Z_{\frac{\alpha}{2}} * \frac{\sigma}{\sqrt{n}}$$

$$Interval\ Estimate: 6.75 \pm 1.96 * \frac{0.75}{\sqrt{60}}$$

$$\boldsymbol{Interval\ Estimate}: \mathbf{6.75\ \pm 0.189}$$

23. What is the LCL for a p-chart when the average daily inspection quantity is 125, and the historical percentage of defectives is 0.10?

- 0.00
- **0.02**
- 0.10
- 0.18

$$\boldsymbol{LCL_{\overline{p}}\ = \overline{p} - 3\sqrt{\frac{\overline{p}(1-\overline{p})}{\overline{n}}}}$$

$$\boldsymbol{LCL_{\overline{p}}\ = 0.10 - 3\sqrt{\frac{0.10(1-0.10)}{125}} = 0.10 - 3\sqrt{0.00072} = 0.019 = 0.02}$$

24. Steph Curry shoots 3-pointers at a success rate of 42%. If he were to take 4 shots in a row, what is the likelihood that he makes all 4?

- 42%
- 12%
- 6%
- **3%**

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

$$P(x = 4) = \binom{4}{4} .42^4 (1 - .42)^{4-4}$$

$$\binom{4}{4} = \frac{4!}{4!\,(4-4)!} = \frac{4!}{4! * 0!} = 1$$

$$\mathrm{P(x = 4)} = (1).42^4(.58)^0 = \mathbf{0.0311\ or\ 3.11\%}$$

25. You manufacture a widget and use an x-bar and R chart to monitor your process, where you sample 3 units in each subgroup, and R-bar = 16.0. Estimate the population standard deviation for this process.

- 16.0
- **9.5**
- 27.1
- 13.2

We divide R-bar by the factor d_2, which is based on the n=3 sample size.

$$\textit{Population Standard Deviation} = \hat{\sigma} = \frac{\bar{R}}{d_2} = \frac{16}{1.693} = 9.5$$