1. A random sample of size 25 from an \( N(\mu, \sigma^2) \) distribution has mean \( \bar{x} = 14.3 \). Find the margin of error \( E \) of the 85% confidence interval \([\bar{x} - E, \bar{x} + E]\) for \( \mu \).

(A) 3.56  
(B) 3.12  
(C) 2.79  
(D) 2.13  
(E) 1.98  
(F) 1.87  
(G) 1.80  
(H) 1.73  
(I) 1.52  
(J) 0.89

Solution: The \( z \)-critical point is \( z_{a/2} = z_{0.075} = 1.44 \). So the margin of error is \( E = z_{a/2} \frac{\sigma}{\sqrt{n}} = 1.44 \times 6/5 = 1.728 \).
2. Determine whether each of the following statements is true (T) or false (F):

(a) In the statement
“A shipment of 1000 fuses contains 3 defectives. A sample of 25 fuses contained 3 defectives”
the highlighted number is a statistic.
(b) Let $X_1, X_2$ be i.i.d. observations from a distribution with mean $\mu$, and consider the estimators:

$$\hat{\mu}_1 = 0.5X_1 + 0.5X_2 \text{ and } \hat{\mu}_2 = 0.3X_1 + 0.7X_2$$

of $\mu$. Then $\hat{\mu}_1$ is an unbiased estimator, whereas $\hat{\mu}_2$ is biased.
(c) Let $X_1, X_2$ be i.i.d. observations from a distribution with mean $\mu$, and consider the estimators:

$$\hat{\mu}_1 = 0.4X_1 + 0.6X_2 \text{ and } \hat{\mu}_2 = 0.7X_1 + 0.3X_2$$

of $\mu$. Then $\hat{\mu}_1$ is a better estimator than $\hat{\mu}_2$, since it has smaller variance.
(d) Let $X_1, X_2, \ldots, X_n$ be a random sample from a $U[0, \theta]$ distribution. Then $(X_{\text{max}} + X_{\text{min}})/2$ is an unbiased estimator of $\theta/2$.

The answers are, respectively,

(A) (F, F, T, T)
(B) (F, T, F, T)
(C) (F, F, T, F)
(D) (F, F, T, F)
(E) (F, T, T, T)
(F) (F, T, T, F)
(G) (F, T, T, T)
(H) (T, F, F, T)
(I) (T, F, F, F)
(J) (T, F, F, F)

Solution: (a) F: The number 3 is associated to the population (rather than the sample); it is a population parameter and not a sample statistic.

(b) F: Both estimators have expected value equal to $\mu$. Therefore, they are both unbiased.

(c) T: $\hat{\mu}_1$ is indeed better. The ratio of the variances of $\hat{\mu}_1$ over the variance of $\hat{\mu}_2$ is

$$\frac{(0.4)^2 + (0.6)^2}{(0.7)^2 + (0.3)^2} = 0.537.$$ 

So $\hat{\mu}_1$ has smaller variance.

(d) T: This is indeed true. It involves a few calculations using order statistics, which came up in the suggested problems from the statistics textbook. I’m assuming here simply that you remember the result of those calculations.
3. A random sample of 4 coffee cans is taken from a production line and the contents weighed. The weights (in oz.) are as follows: 26.5, 25.0, 27.0, 25.5.

Find the standard error of the mean (SEM).

(A) 0.91  
(B) 0.83  
(C) 0.57  
(D) 0.46  
(E) 0.41  
(F) 0.39  
(G) 0.34  
(H) 0.30  
(D) 0.23  
(J) 0.15

Solution:

- Sample mean: \( \bar{x} = 26.0 \)
- Sample variance: 
  \[
  s^2 = \frac{1}{n-1} \sum (x_i - \bar{x})^2 = \left(\frac{(26.5-26.0)^2 + (25.0-26.0)^2 + (27.0-26.0)^2 + (25.5-26.0)^2}{3}\right) = 0.833
  \]
- SEM: 
  \[
  s/\sqrt{n} = 0.9128/2 = 0.456.
  \]
4. Let $X$ be a Bernoulli r.v. with $P(X = 1) = p$ and $P(X = 0) = 1 - p$. We want to test

$$H_0 : p = 1/5 \text{ vs. } H_1 : p = 4/5.$$ 

Suppose that based on two i.i.d. observations $X_1$ and $X_2$ we adopt the following decision rule:

$$\begin{align*}
\text{do not reject } H_0 & \quad \text{if } X_1 + X_2 = 0 \text{ or } 1 \\
\text{reject } H_0 & \quad \text{if } X_1 + X_2 = 2.
\end{align*}$$

The probabilities of type I and type II errors are, respectively

(A) $\alpha = 0.26, \beta = 0.82$
(B) $\alpha = 0.82, \beta = 0.26$
(C) $\alpha = 0.36, \beta = 0.04$
(D) $\alpha = 0.04, \beta = 0.36$
(E) $\alpha = 0.05, \beta = 0.25$
(F) $\alpha = 0.25, \beta = 0.05$
(G) $\alpha = 0.09, \beta = 0.73$
(H) $\alpha = 0.73, \beta = 0.09$
(I) $\alpha = 0.27, \beta = 0.06$
(J) $\alpha = 0.06, \beta = 0.27$

Solution: The probability of a type I error is

$$\alpha = P(\text{reject } H_0 | H_0) = P(X_1 + X_2 = 2 | p = 1/5) = P(X_1 = 1 | p = 1/5)P(X_2 = 1 | p = 1/5) = (1/5)^2 = 0.04.$$ 

The probability of a type II error is

$$\beta = P(\text{not reject } H_0 | H_1) = 1 - P(\text{reject } H_0 | H_1) = 1 - P(X_1 = 1 | p = 4/5)P(X_2 = 1 | p = 4/5) = 1 - (4/5)^2 = 0.36.$$
5. An EPA researcher wants to design a study to estimate the mean lead level of fish in a lake located near an industrial area. Based on past sample data, the researcher estimates that $\sigma$ for the lead level in the fish population is approximately 0.02 mg/g. He wants to use a 90% confidence interval having a margin of error no greater than 0.01 mg/g. How many fish does he need to catch?

\[
(A) \quad 11 \\
(B) \quad 39 \\
(C) \quad 33 \\
(D) \quad 37 \\
(E) \quad 42 \\
(F) \quad 49 \\
(G) \quad 56 \\
(H) \quad 65 \\
(I) \quad 78 \\
(J) \quad 104
\]

**Solution:** We have:

- $\sigma = 0.02$
- $z_{\alpha/2} = z_{0.05} = 1.64$
- $E = 0.01$

The sample size necessary for a confidence level $\alpha = 0.1$ and margin of error $E = 0.01$ is

\[
 n = \left[ \frac{z_{\alpha/2} \sigma}{E} \right]^2 = \left[ \frac{1.64 \times 0.02}{0.01} \right]^2 = 10.76.
\]

We round it up to 11 fish.
6. A random sample of size 12 is drawn from a normal distribution. The mean of the sample is 67 and \( s = 2.5 \). Calculate the margin of error \( E \) of a 90% \( t \)-interval \([\bar{x} - E, \bar{x} + E]\) for the population mean \( \mu \).

(A) 1.60
(B) 2.50
(C) 0.20
(D) 3.10
(E) 1.30
(F) 1.70
(G) 2.20
(H) 0.60
(I) 0.80
(J) 1.80

Solution: Summarizing the information:

- \( n = 12 \)
- \( s = 2.5 \)
- \( t_{n-1,\alpha/2} = t_{11,0.05} = 1.796 \)

The margin of error is

\[
E = t_{n-1,\alpha/2} \frac{s}{\sqrt{n}} = 1.796 \times \frac{2.5}{\sqrt{12}} = 1.30
\]
7. This and the next two problems refer to the following situation. A bottling company uses a filling machine to fill bottles. A bottle is to contain 400 milliliters of beverage. The actual amount is normally distributed with a standard deviation of 1.0 ml. The purchase of a new machine is contemplated. Based on a sample of 14 bottles filled by the new machine, the sample mean is 410 ml. and the standard deviation is $s = 0.5$. At a confidence level 0.95, does the new machine have significantly less variation than the current machine? To answer this question, we will test the hypotheses $H_0 : \sigma \geq 1.0$ vs. $H_1 : \sigma < 1.0$.

What is the relevant statistic for this test, and what is its value?

(A) $\chi^2 = 7.35$
(B) $\chi^2 = 3.25$
(C) $\chi^2 = 2.13$
(D) $\chi^2 = 4.15$
(E) $\chi^2 = 3.89$
(F) $t = 7.83$
(G) $t = 1.7$
(H) $t = 7.35$
(I) $t = 8.50$
(J) $t = 1.05$

Solution: The main information is summarized here:

- Sample size $n = 14$
- $\sigma_0 = 1.0$
- $s = 0.5$

The test statistic is

$$
\chi^2 = \frac{(n - 1)s^2}{\sigma_0^2} = \frac{13 \times (0.5)^2}{(1.0)^2} = 3.25.
$$
8. (Continuation of the previous problem.) Assume the same information given in problem 7. What is the value of the critical point at level \( \alpha = 0.05 \) needed for this test, and do we have enough evidence to reject \( H_0 \)?

(A) \( \chi^2_{13,0.05} = 22.36 \); we reject \( H_0 \)

(B) \( \chi^2_{13,0.05} = 22.36 \); we do not reject \( H_0 \)

(C) \( \chi^2_{13,0.95} = 5.89 \); we reject \( H_0 \)

(D) \( \chi^2_{13,0.95} = 5.89 \); we do not reject \( H_0 \)

(E) \( \chi^2_{14,0.05} = 23.68 \); we reject \( H_0 \)

(F) \( \chi^2_{14,0.05} = 23.68 \); we do not reject \( H_0 \)

(G) \( t_{13,0.05} = 3.012 \); we reject \( H_0 \)

(H) \( t_{13,0.05} = 3.012 \); we do not reject \( H_0 \)

(I) \( t_{14,0.05} = 2.977 \); we reject \( H_0 \)

(J) \( t_{14,0.05} = 2.977 \); we do not reject \( H_0 \)

Solution: The critical point is \( \chi^2_{n-1,1-\alpha} = \chi^2_{13,0.95} = 5.89 \). Since \( \chi^2 = 3.25 \) is less than \( \chi^2_{13,0.95} \), we are justified in rejecting \( H_0 \) at the confidence level \( \alpha = 0.95 \).
9. (Continuation of the previous problem.) Assume the same information given in problem 7. Knowing that the P-value of the test is 0.003, determine whether the following statements are true (T) or false (F).

(a) Based on this P-value, the null-hypothesis can be rejected at the confidence level 0.99.
(b) \( P(\chi^2_{13} > \chi^2) = 0.003 \), where \( \chi^2 \) is the value of the correct test statistic for this problem.
(c) \( P(T_{13} < t) = 0.003 \), where \( t \) is the value of the correct test statistic for this problem.

These statements are, respectively

(A)  T, F, T
(B)  E, T, T
(C)  T, E, F
(D)  E, T, F
(E)  E, F, T
(F)  E, F, F
(G)  T, T, F
(H)  T, T, T

Solution: (a) F: We reject the null hypothesis since the P-value is smaller than the \( \alpha = 0.01 \) level.
(b) F: The P-value for this test is \( P(\chi^2_{13} \leq \chi^2) = 0.003 \). Note the direction of the inequality.
(c) F: The correct test statistic is \( \chi^2 \), and not \( t \).
10. Production lines in a manufacturing plant are set to make steel ball bearings with a diameter of 1 micron. Ten ball bearings were randomly selected from two production lines and the following QQ-plot of the samples was obtained, where the straight line is a 45° line through the origin:

Which of the following is correct:

(A) Experiment has independent samples design; second line tends to produce balls of larger diameter.
(B) Experiment has matched pairs design; second line tends to produce balls of larger diameter.
(C) Experiment has matched pairs design; first line tends to produce balls of larger diameter.
(D) Experiment has independent samples design; first line tends to produce balls of larger diameter.

Solution: The experiment has independent samples design. Most of the points lie under the 45° line, so the first line quantiles tend to be greater than the corresponding second line quantiles. This indicates that the first production line tends to produce balls of larger diameter.
11. This and the next problem refer to the following situation. Two brands of water filters are to be compared in terms of the mean reduction in impurities measured in parts per million. Fifteen water samples were tested with each filter and reduction in the impurity level was measured, resulting in the following data:

Filter 1: \( n_1 = 15 \) \( \bar{x} = 8.0 \) \( s_1^2 = 4.5 \)

Filter 2: \( n_2 = 15 \) \( \bar{y} = 6.5 \) \( s_2^2 = 2.0 \)

Calculate a 95% confidence interval for the mean difference \( \mu_1 - \mu_2 \) between the two filters, assuming \( \sigma_1^2 = \sigma_2^2 \).

Is there a statistically significant difference at \( \alpha = 0.05 \) between the two filters? Consider here a two-sided test:

\[ H_0: \mu_1 = \mu_2 \text{ vs. } H_1: \mu_1 \neq \mu_2. \]

(A) CI: \([-0.15, 2.85]\); we reject \( H_0 \)

(B) CI: \([-0.15, 2.85]\); we do not reject \( H_0 \)

[C] CI: \([0.15, 2.85]\); we reject \( H_0 \)

(D) CI: \([-2.85, 0.15]\); we reject \( H_0 \)

(E) CI: \([-2.85, 0.15]\); we do not reject \( H_0 \)

(F) CI: \([0.15, 2.85]\); we do not reject \( H_0 \)

(G) CI: \([-1.85, -0.15]\); we reject \( H_0 \)

(H) CI: \([-1.85, -0.15]\); we do not reject \( H_0 \)

(I) CI: \([1.85, 0.45]\); we reject \( H_0 \)

(J) CI: \([1.85, 0.45]\); we do not reject \( H_0 \)

Solution: The pooled standard deviation is

\[
s = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}} = \sqrt{\frac{14 \times 4.5 + 14 \times 2.0}{28}} = 1.803.
\]

A 95% confidence interval for \( \mu_1 - \mu_2 \) has margin of error:

\[
E = t_{n_1 + n_2 - 2, \alpha/2} s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} = t_{28, 0.025} s \sqrt{\frac{2}{15}} = 2.048 \times 1.803 \times 0.37 = 1.35
\]

Therefore,

\[
\mu_1 - \mu_2 = \bar{x} - \bar{y} \pm E = 1.5 \pm 1.35.
\]

The confidence interval is \([0.15, 2.85]\). As the interval does not contain 0, we reject \( H_0 \) at \( \alpha = 0.05 \).
12. (Continuation of the previous problem.) Assume the same information given in problem 11. We wish to calculate a 95% confidence interval for the mean difference $\mu_1 - \mu_2$ between the two filters, now without assuming $\sigma_1^2 = \sigma_2^2$. This can be done using a $t$-statistic with $\nu$ degrees of freedom, where $\nu$ is estimated from the sample. What is $\nu$, rounded off to the nearest integer?

(A) 24  
(B) 22  
(C) 20  
(D) 18  
(E) 16  
(F) 14  
(G) 12  
(H) 10  
(I) 8  
(J) 6

**Solution:** The degrees of freedom are obtained as follows:

$$w_1 = \frac{s_1^2}{n_1} = \frac{4.5}{15} = 0.30, \quad w_2 = \frac{s_2^2}{n_2} = \frac{2.0}{15} = 0.13, \quad \nu = \frac{(w_1 + w_2)^2}{\frac{w_1^2}{n_1-1} + \frac{w_2^2}{n_2-1}} = \frac{14 \times (0.30 + 0.13)^2}{(0.30)^2 + (0.13)^2} = 24.21.$$  

Rounded off to the nearest integer, $\nu = 24$. 

12
13. This and the next two problem refer to the following situation. To determine whether glaucoma affects the corneal thickness, measurements were made in 8 people affected by glaucoma in one eye but not in the other. Suppose that the sample mean of corneal thickness for the affected eyes was \( \bar{x} = 480 \) and the sample mean of corneal thickness for the unaffected eyes was \( \bar{y} = 485 \). Also suppose that the sample standard deviation for the differences of corneal thickness of the 8 pairs was \( s_d = 2.5 \). We want to test the hypotheses: \( H_0 : \mu_1 = \mu_2 \) vs. \( H_1 : \mu_1 \neq \mu_2 \). Which of the following is correct?

This experiment has:

(A) an independent samples design and the test statistic is \( t = 10.27 \)
(B) a matched pairs design and the test statistic is \( t = 8.05 \)
(C) an independent samples design and the test statistic is \( t = 5.98 \)
(D) an independent samples design and the test statistic is \( t = -5.66 \)
(E) an independent samples design and the test statistic is \( t = 8.05 \)
(F) a matched pairs design and the test statistic is \( t = 5.98 \)
(G) a matched pairs design and the test statistic is \( t = -2.63 \)
(H) an independent samples design and the test statistic is \( t = -2.63 \)
(I) a matched pairs design and the test statistic is \( t = 10.27 \)
(J) a matched pairs design and the test statistic is \( t = -5.66 \)

**Solution:** This experiment has a matched pairs design. The test statistic is

\[
t = \frac{\bar{x} - \bar{y}}{s_d / \sqrt{n}} = \frac{480 - 485}{2.5 / \sqrt{8}} = -5.66.
\]
14. (Continuation of the previous problem.) Assume the same information given in problem 13. Calculate the $t$-critical point to test the hypotheses $H_0: \mu_1 = \mu_2$ vs. $H_1: \mu_1 \neq \mu_2$ at $\alpha = 0.05$. For what values of the $t$-statistics can we reject $H_0$ at this $\alpha$?

(A) $t_{n-1, \alpha/2} = 2.365$, $|t| < t_{n-1, \alpha/2}$

(B) $t_{n-1, \alpha/2} = 2.365$, $t < t_{n-1, \alpha/2}$

(C) $t_{n-1, \alpha/2} = 2.365$, $t > t_{n-1, \alpha/2}$

(D) $t_{n-1, \alpha/2} = 2.365$, $|t| > t_{n-1, \alpha/2}$

(E) $t_{n-1, \alpha/2} = 1.434$, $|t| < t_{n-1, \alpha/2}$

(F) $t_{n-1, \alpha/2} = 1.434$, $t < t_{n-1, \alpha/2}$

(G) $t_{n-1, \alpha/2} = 1.434$, $t > t_{n-1, \alpha/2}$

(H) $t_{n-1, \alpha/2} = 1.434$, $|t| > t_{n-1, \alpha/2}$

**Solution:** The critical point of the $t$-statistic at $\alpha = 0.05$ is

$$t_{n-1, \alpha/2} = t_{7, 0.025} = 2.365.$$