Answers - Statistical inference

Year: University Year 1

Answer to Question 1

a. $\bar{x} = \dfrac{116}{8} = \mathbf{14.5}$.

b. Deviations: $-2.5,\ 0.5,\ -3.5,\ 3.5,\ -0.5,\ 1.5,\ -1.5,\ 2.5$; squares add to 42. So $s^2 = \dfrac{42}{7} = \mathbf{6}$ and $s = \sqrt{6} \approx \mathbf{2.449}$.

c. $\dfrac{s}{\sqrt{n}} = \dfrac{2.449}{\sqrt{8}} \approx \mathbf{0.866}$.

d. The deviations are measured from $\bar{x}$, not from the unknown μ, which makes them slightly too small on average. Dividing by $n - 1$ makes s^2 an unbiased estimator of σ^2.

Answer to Question 2

a. The sample size $n = 25$ is reasonably large, so by the central limit theorem $\bar{X}$ is approximately normal whatever the shape of the population. Its mean is $\mu = 40$ and its variance is $\dfrac{\sigma^2}{n} = \dfrac{100}{25} = 4$, so $\bar{X} \approx N(40, 4)$, with standard error $\sqrt{4} = 2$.

b. Standardise: $z = \dfrac{43 - 40}{2} = 1.5$, so $P(\bar{X} > 43) = 1 - \Phi(1.5) = \mathbf{0.0668}$.

c. The limits are one standard error either side of the mean: $P(38 < \bar{X} < 42) = P(-1 < Z < 1) = 2\Phi(1) - 1 = \mathbf{0.6827}$.

d. The standard error is $\dfrac{10}{\sqrt{n}}$. We need $\dfrac{10}{\sqrt{n}} \leq 1$, so $\sqrt{n} \geq 10$ and $n \geq \mathbf{100}$. Halving the standard error needs four times as many observations.

Answer to Question 3

a. With σ known, the interval is $\bar{x} \pm 1.96\dfrac{\sigma}{\sqrt{n}}$. The standard error is $\dfrac{6}{\sqrt{36}} = 1$, so $72.4 \pm 1.96 \times 1 = 72.4 \pm 1.96$, giving $(70.44,\ 74.36)$.

b. Here σ is unknown and n is small, so use the t distribution with $n - 1 = 9$ degrees of freedom: $t_{9,\,0.975} = 2.262$. The estimated standard error is $\dfrac{2.1}{\sqrt{10}} = 0.664$. Then $15.2 \pm 2.262 \times 0.664 = 15.2 \pm 1.50$, giving $(13.70,\ 16.70)$. The t value is larger than 1.96, which widens the interval to allow for estimating σ.

c. The sample proportion is $\hat{p} = \frac{84}{200} = 0.42$, with $SE = \sqrt{\frac{0.42 \times 0.58}{200}} = 0.0349$. Then $0.42 \pm 1.96 \times 0.0349 = 0.42 \pm 0.068$, giving $(0.352,\ 0.488)$.

Answer to Question 4

a. With σ known, the test statistic is $z = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{509 - 500}{20/4} = 1.8$. For a two-sided alternative, $p = 2(1 - \Phi(1.8)) = 0.072$. Since $0.072 > 0.05$, we **do not reject** H_0: there is not enough evidence that the mean differs from 500.

b. With $H_1 : \mu > 500$, only the upper tail counts: $p = 1 - \Phi(1.8) = 0.036 < 0.05$, so we **reject** H_0. The same data give different conclusions, which is why the alternative hypothesis must be chosen before seeing the data.

c. σ is unknown, so use $t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = \frac{0.6}{0.9/\sqrt{12}} = 2.309$ with 11 degrees of freedom. The two-sided critical value is $t_{11,\,0.975} = 2.201$. Since $2.309 > 2.201$, we **reject** H_0 ($p \approx 0.041$): the mean appears to differ from 4.

Answer to Question 5

a. H_0: the die is fair, so each score has probability $\frac{1}{6}$; H_1: the die is not fair. Under H_0 the expected frequency of each score is $\frac{120}{6} = 20$.

b. The differences $O - E$ are $5, -3, -5, 3, 4, -4$, so the values of $(O - E)^2$ are $25, 9, 25, 9, 16, 16$, with total 100. Since every $E = 20$, $\chi^2 = \frac{100}{20} = \mathbf{5.0}$.

c. There are 6 categories and one constraint (the total), so there are $6 - 1 = 5$ degrees of freedom. The 5% critical value of χ^2_5 is 11.07. Since $5.0 < 11.07$, we do not reject H_0: there is **no evidence** that the die is biased.

Answer to Question 6

a. The means are $\bar{x} = 3$ and $\bar{y} = \frac{31}{5} = 6.2$. Using $S_{xx} = \sum(x - \bar{x})^2$ and similar formulas: $S_{xx} = 4 + 1 + 0 + 1 + 4 = 10$;
$S_{xy} = (-2)(-3.2) + (-1)(-1.2) + 0 + (1)(1.8) + (2)(2.8) = 15$;
$S_{yy} = 10.24 + 1.44 + 0.04 + 3.24 + 7.84 = 22.8$.

b. The least-squares gradient is $b = \frac{S_{xy}}{S_{xx}} = \frac{15}{10} = 1.5$. The line passes through $(\bar{x}, \bar{y})$, so $a = \bar{y} - b\bar{x} = 6.2 - 4.5 = 1.7$. The line is $y = 1.7 + 1.5x$.

c. $r = \frac{S_{xy}}{\sqrt{S_{xx} S_{yy}}} = \frac{15}{\sqrt{10 \times 22.8}} \approx \mathbf{0.993}$. This is very close to 1: a very strong

positive linear relationship, so the line fits the data well.

Answer to Question 7

a. $SE = \dfrac{2.5}{\sqrt{40}} = 0.395$; $18.6 \pm 1.96 \times 0.395 = 18.6 \pm 0.77$: $(17.83,\ 19.37)$ hours (n is large, so z is used).
b. No: 20 lies above the interval, so the data suggest that the true mean is lower than advertised.
c. $18.6 \pm 2.576 \times 0.395$: $(17.58,\ 19.62)$. More confidence requires a wider net.
d. $1.96 \times \dfrac{2.5}{\sqrt{n}} \leq 0.5 \iff n \geq 96.04$: **97 phones**.

Answer to Question 8

a. $H_0 : \mu = 250$ against $H_1 : \mu < 250$, where μ is the mean volume poured. The test is one-sided because the concern is short measures.
b. The standard error is $\dfrac{8}{\sqrt{25}} = 1.6$, so $z = \dfrac{246.5 - 250}{1.6} \approx -2.19$. The p-value is $\Phi(-2.19) \approx$ **0.014**.
c. At the 5% level, $0.014 < 0.05$, so we **reject** H_0: there is evidence that the machine gives short measures. At the 1% level, $0.014 > 0.01$, so we **do not reject** H_0. The conclusion depends on how strong the evidence must be.
d. A Type I error means rejecting a true H_0: concluding that the machine pours short measures, and paying for a repair, when in fact it is correctly set.

Answer to Question 9

a. $\hat{p} = \dfrac{520}{1000} = 0.52$ and $SE = \sqrt{\dfrac{0.52 \times 0.48}{1000}} = 0.0158$. The margin of error is $1.96 \times 0.0158 = 0.031$, so the interval is 0.52 ± 0.031, that is $(0.489,\ 0.551)$.
b. Under H_0, $p = 0.5$, so $SE = \sqrt{\dfrac{0.5 \times 0.5}{1000}} = 0.0158$ and $z = \dfrac{0.52 - 0.5}{0.0158} = 1.265$. The one-sided p-value is about $0.103 > 0.05$, so we **do not reject** H_0.
c. No. The confidence interval includes values below 0.5, and the test is not significant, so the poll does not establish a majority. The headline overstates the evidence.
d. The margin of error is largest when $p = 0.5$. We need $1.96\sqrt{\dfrac{0.25}{n}} \leq 0.02$, so $n \geq \left(\dfrac{1.96 \times 0.5}{0.02}\right)^2 = 2401$: about **2 400 voters**.

Answer to Question 10

a. The two scores belong to the same student, so they are not independent; working with the differences removes the variation between students.

b. $\bar{d} = \dfrac{28}{8} = \mathbf{3.5}$; squared deviations sum to 46, so $s_d^2 = \dfrac{46}{7}$, $s_d \approx \mathbf{2.56}$.

c. $H_0 : \mu_d = 0$, $H_1 : \mu_d > 0$. $t = \dfrac{3.5}{2.56/\sqrt{8}} \approx \mathbf{3.86}$ with 7 df; critical $t_{7,\,0.95} = 1.895$. **Reject** H_0: evidence of improvement.

d. There is no control group: scores might have improved anyway because of normal revision or a practice effect.

Answer to Question 11

a. $H_0 : \mu_1 = \mu_2$ (the lines have the same mean length) against $H_1 : \mu_1 \neq \mu_2$, a two-sided test because no direction is suggested.

b. For independent samples the variances add:
$SE = \sqrt{\dfrac{s_1^2}{n_1} + \dfrac{s_2^2}{n_2}} = \sqrt{\dfrac{1.5^2}{50} + \dfrac{1.8^2}{60}} = \sqrt{0.045 + 0.054} = \mathbf{0.315}$. The samples are large, so the sample standard deviations can be used in place of σ.

c. $z = \dfrac{12.4 - 11.8}{0.315} = \dfrac{0.6}{0.315} = \mathbf{1.91}$, and $p = 2(1 - \Phi(1.91)) \approx \mathbf{0.057}$.

d. At 5%, $0.057 > 0.05$: not significant, so we do not reject H_0. At 10%, $0.057 < 0.10$: significant, so we reject H_0. The evidence is borderline; more data would help.

Answer to Question 12

a. H_0: accidents are uniformly distributed over the five days; H_1: they are not.

b. Each expected frequency is 26. $\chi^2 = \dfrac{16 + 16 + 64 + 1 + 81}{26} = \dfrac{178}{26} \approx \mathbf{6.85}$.

c. 4 degrees of freedom, critical value 9.488. 6.85: **do not reject** H_0.

d. Friday contributes $\dfrac{81}{26} \approx 3.1$: possibly more traffic and tiredness at the end of the week, though the overall evidence is not significant.

Answer to Question 13

a. $\bar{x} = 21.33$, $\bar{y} = 189.17$. $b = \dfrac{1586.67}{167.33} = 9.48$, $a = 189.17 - 9.48 \times 21.33 = -13.1$: $y = -13.1 + 9.48x$.

b. Each extra degree is associated with about 9.5 more ice creams sold.

c. $-13.1 + 9.48 \times 25 \approx$ **224 ice creams**.
d. $r = \dfrac{1586.67}{\sqrt{167.33 \times 15120.8}} \approx$ **0.997**.
e. No: 5 °C is far outside the data range (extrapolation), and the line would predict only about 34 sales, with no evidence that the relationship stays linear.

Answer to Question 14

a. The standard error is $\dfrac{4}{\sqrt{16}} = 1$. For a one-sided test at 5%, the critical value is $z = -1.645$, so H_0 is rejected when $\bar{x} < 100 - 1.645 \times 1 =$ **98.36** (more precisely 98.355).
b. A Type I error is rejecting H_0 when it is true. The critical value was chosen so that this probability equals the significance level: **0.05**.
c. A Type II error is failing to reject H_0 when it is false. If $\mu = 97$, then $P(\bar{X} \geq 98.355) = P\left(Z \geq \dfrac{98.355 - 97}{1}\right) = P(Z \geq 1.355) \approx$ **0.088**.
d. Power $= 1 - P(\text{Type II}) = 1 - 0.088 =$ **0.912**: the test detects a 3 g shortfall about 91% of the time. Taking larger samples reduces the standard error and increases the power.

Answer to Question 15

a. H_0: sleep quality and exercise level are independent; H_1: they are associated.
b. $E = \dfrac{100 \times 50}{170} = 29.41$. Others: $35.29,\ 35.29$ (good) and $20.59,\ 24.71,\ 24.71$ (poor).
c. $\chi^2 = \sum \dfrac{(O - E)^2}{E} \approx$ **13.80**.
d. $(2 - 1)(3 - 1) = 2$ degrees of freedom; the 1% critical value is 9.210. $13.80 > 9.210$: **reject** H_0. Poor sleep is over-represented in the low-exercise group.
e. The test shows association, not causation: people who sleep badly may exercise less, or both may depend on age or health.

Answer to Question 16

a. $E[X^2] = \dfrac{1 + 4 + 9 + 16 + 25 + 36}{6} = \dfrac{91}{6}$, so $\sigma^2 = \dfrac{91}{6} - \dfrac{49}{4} = \dfrac{35}{12}$.
b. $\bar{X} \approx N\left(3.5, \dfrac{35}{360}\right)$, standard error 0.312.

c. $P(\bar{X} > 4) = 1 - \Phi\left(\dfrac{0.5}{0.312}\right) = 1 - \Phi(1.60) \approx$ **0.054**; expected $200 \times 0.054 \approx$ **11** samples.
d. By symmetry about 11 below 3 and about 89 in each middle class; observed $11, 91, 86, 12$: very close.
e. The central limit theorem: the mean of many independent variables with finite variance is approximately normal, whatever their original distribution.

Answer to Question 17

a. $\bar{x} = \dfrac{304.2}{6} =$ **50.70** kN; $s \approx$ **1.185** kN.
b. $SE = \dfrac{1.185}{\sqrt{6}} = 0.484$; $t_{5,\,0.975} = 2.571$. 50.70 ± 1.24: $(49.46,\ 51.94)$.
c. $t = \dfrac{0.70}{0.484} =$ **1.45**; critical value $t_{5,\,0.95} = 2.015$. **Do not reject** H_0.
d. No: the data do not show that the mean exceeds 50 kN. Failing to reject H_0 does not prove anything either. Test more cables before deciding.

Answer to Question 18

a. $t = 0.72\sqrt{\dfrac{8}{1 - 0.5184}} = 0.72 \times 4.076 \approx$ **2.93**.
b. $t_{8,\,0.975} = 2.306$; $2.93 > 2.306$: **reject** H_0. There is a significant positive correlation.
c. Correlation does not imply causation: neither variable causes the other.
d. Hot weather increases both ice cream sales and the number of swimmers. Record the temperature and the number of swimmers, and study rescues per swimmer.

Answer to Question 19

a. $SE = \dfrac{8}{\sqrt{n}}$; reject when $|\bar{x}| > 1.96\,SE$.
b. $SE = 0.8$; reject when $\bar{x} > 1.568$. Power $\approx P\left(Z > \dfrac{1.568 - 2}{0.8}\right) = P(Z > -0.54) \approx$ **0.705**.
c. We need $\dfrac{2}{8/\sqrt{n}} \geq 1.96 + 1.2816$, giving $n \geq 168.1$: **169 people**.
d. A smaller significance level or higher power requires a larger sample, which costs more; too small a trial may miss a real effect.

Answer to Question 20

a. By linearity, $E[\bar{X}] = \frac{1}{n}\sum E[X_i] = \frac{1}{n}(n\mu) = \mu$. Since the X_i are independent, their variances add: $\mathrm{Var}(\bar{X}) = \frac{1}{n^2}\sum \mathrm{Var}(X_i) = \frac{1}{n^2}(n\sigma^2) = \dfrac{\sigma^2}{n}$.

b. Expand the square: $\sum (X_i - \bar{X})^2 = \sum X_i^2 - 2\bar{X}\sum X_i + n\bar{X}^2$. Since $\sum X_i = n\bar{X}$, this is $\sum X_i^2 - 2n\bar{X}^2 + n\bar{X}^2 = \sum X_i^2 - n\bar{X}^2$.

c. Take expectations:
$E\left[\sum (X_i - \bar{X})^2\right] = n(\sigma^2 + \mu^2) - n\left(\dfrac{\sigma^2}{n} + \mu^2\right) = n\sigma^2 - \sigma^2 = (n-1)\sigma^2$.

d. Dividing by $n-1$ gives $E[S^2] = \sigma^2$, so S^2 is unbiased. With $\frac{1}{n}$ instead, the expectation is $\dfrac{n-1}{n}\sigma^2$, so the bias is $-\dfrac{\sigma^2}{n}$: it underestimates σ^2, but only slightly for large samples.