Solutions to Probability and Statistics Problems

Question 1

(a) Stem and Leaf Display

The given data appears to have some formatting issues (A1, A2, etc. mixed with numbers). Assuming the numerical values are:

31, 35, 36, 36, 37, 38, 40, 40, 54, 55, 58, 62, 66, 66, 67, 68, 75

Stem-and-Leaf Display:

```
3 | 1 5 6 6 7 8
4 | 0 0
5 | 4 5 8
6 | 2 6 6 7 8
7 | 5
```

Interesting Features:

- The data is bimodal with clusters around 30s and 60s
- There's a gap between 40 and 54
- The distribution is not symmetric

(b) Median and Quartiles

Data: 16, 20, 26, 30, 35, 39, 45, 46, 48, 50 (n=10)

Median (Q2): $(5\text{th} + 6\text{th})/2 = (35 + 39)/2 = 37$

Lower Fourth (Q1): Median of first half = 26

Upper Fourth (Q3): Median of second half = 46

(c) Oxidation-Induction Time

Data: 87, 103, 105, 130, 132, 145, 160, 180, 195, 211

(i) Sample Variance and Standard Deviation

Mean (μ) = $(87 + 103 + \cdots + 211)/10 = 144.8$

Variance (s^2) = $\frac{\sum(x_i-\mu)^2}{n-1} = \frac{(87-144.8)^2+\cdots+(211-144.8)^2}{9} = 1566.4$

Standard Deviation (s) = $\sqrt{1566.4} \approx 39.58$ minutes

(ii) Conversion to Hours
Variance in hours $= \frac{1566.4}{60^2} \approx 0.4351$ hours^2
Standard Deviation in hours $= \frac{39.58}{60} \approx 0.6597$ hours

Question 2

(a) Probability Proofs

First Proof: $P(A \cap B') = P(A) - P(A \cap B)$

This follows from the fact that A can be written as the union of two disjoint events: $A = (A \cap B) \cup (A \cap B')$ Thus $P(A) = P(A \cap B) + P(A \cap B')$

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

This follows from the inclusion-exclusion principle, as the intersection is counted twice in $P(A) + P(B)$.

(b) Aircraft Discovery Probability

Let:

- D = discovered ($P(D) = 0.7$)
- D' = not discovered ($P(D') = 0.3$)
- L = has locator

Given:

- $P(L|D) = 0.6 \Rightarrow P(L'|D) = 0.4$
- $P(L'|D') = 0.9 \Rightarrow P(L|D') = 0.1$

We need $P(D'|L)$:
$P(L) = P(L|D)P(D) + P(L|D')P(D') = (0.6)(0.7) + (0.1)(0.3) = 0.45$
$P(D'|L) = \frac{P(L|D')P(D')}{P(L)} = \frac{(0.1)(0.3)}{0.45} \approx 0.0667$ or 6.67%

(c) Bayes' Theorem Application

Let:

- D = day visit ($P(D) = 0.2$)
- N = one-night visit ($P(N) = 0.5$)
- T = two-night visit ($P(T) = 0.3$)
- B = makes purchase

Given:

- $P(B|D) = 0.1$
- $P(B|N) = 0.3$
- $P(B|T) = 0.2$

$P(B) = P(B|D)P(D) + P(B|N)P(N) + P(B|T)P(T) = (0.1)(0.2) + (0.3)(0.5) + (0.2)(0.3) = 0.23$

$P(D|B) = \frac{P(B|D)P(D)}{P(B)} = \frac{(0.1)(0.2)}{0.23} \approx 0.087$ or 8.7%

Question 3

(a) Blood Donor Typing

(i) PMF of Y

Possible values: $Y = 1, 2, 3, 4, 5$
$P(Y = 1) = \frac{2}{5}$ (either a or b first)
$P(Y = 2) = \frac{3}{5} \times \frac{2}{4} = \frac{3}{10}$
$P(Y = 3) = \frac{3}{5} \times \frac{2}{4} \times \frac{2}{3} = \frac{1}{5}$
$P(Y = 4) = \frac{3}{5} \times \frac{2}{4} \times \frac{1}{3} \times \frac{2}{2} = \frac{1}{10}$
$P(Y = 5) = 0$ (must find by 4th trial)

(ii) Graphs

Line graph and histogram would show:

- Y=1: 0.4
- Y=2: 0.3
- Y=3: 0.2
- Y=4: 0.1
- Y=5: 0

(b) Job Candidate Ranking

(i) PMF of X

$P(X = k) = \frac{1}{n}$ for $k = 1, 2, \ldots, n$

(ii) E(X) and V(X)

$E(X) = \frac{n+1}{2}$
$V(X) = \frac{n^2-1}{12}$

(c) Variance Proof

For $V(aX+b)$:

$$\begin{aligned}
&= E[(aX+b-E[aX+b])^2] \\
&= E[(aX+b-a\mu-b)^2] \\
&= E[a^2(X-\mu)^2] \\
&= a^2E[(X-\mu)^2] \\
&= a^2V(X)
\end{aligned}$$

For σ_{aX+b}:

$$= \sqrt{V(aX+b)} = \sqrt{a^2V(X)} = |a|\sigma_X$$

Question 4

(a) Gravel Sales Distribution

(i) CDF

$F(x) = \int_0^x \frac{3}{2}(1-t^2)dt = \frac{3}{2}(x - \frac{x^3}{3})$ for $0 \le x \le 1$

$= 0$ for $x < 0$

$= 1$ for $x > 1$

(ii) E(X)

$E(X) = \int_0^1 x\frac{3}{2}(1-x^2)dx = \frac{3}{2}\int_0^1 (x-x^3)dx = \frac{3}{2}[\frac{1}{2} - \frac{1}{4}] = \frac{3}{8}$

(iii) V(X)

$E(X^2) = \int_0^1 x^2\frac{3}{2}(1-x^2)dx = \frac{3}{2}[\frac{1}{3} - \frac{1}{5}] = \frac{1}{5}$

$V(X) = E(X^2) - [E(X)]^2 = \frac{1}{5} - (\frac{3}{8})^2 = \frac{19}{320} \approx 0.0594$

(iv) σ_X

$\sigma_X = \sqrt{V(X)} \approx \sqrt{0.0594} \approx 0.244$

(b) Reaction Time

$\mu = 1.25$, $\sigma = 0.46$

(i) P(1.00 ¡ X ¡ 1.75)

$Z_1 = \frac{1.00-1.25}{0.46} \approx -0.5435 \Rightarrow P(Z < -0.5435) \approx 0.2934$

$Z_2 = \frac{1.75-1.25}{0.46} \approx 1.0870 \Rightarrow P(Z < 1.0870) \approx 0.8616$

$P = 0.8616 - 0.2934 \approx 0.5682$ or 56.82%

(ii) P(X ¿ 2)

$Z = \frac{2-1.25}{0.46} \approx 1.6304$

$P(Z > 1.6304) \approx 1 - 0.9486 = 0.0514$ or 5.14%

(c) Binomial Distribution Proofs

(i) E[X] = np

$$
\begin{aligned}
E[X] &= \sum_{k=0}^{n} k \binom{n}{k} p^k (1-p)^{n-k} \\
&= n \sum_{k=1}^{n} \binom{n-1}{k-1} p^k (1-p)^{n-k} \\
&= np \sum_{j=0}^{n-1} \binom{n-1}{j} p^j (1-p)^{n-1-j} \quad [\text{let } j = k-1] \\
&= np(p + (1-p))^{n-1} = np
\end{aligned}
$$

(ii) V[X] = np(1-p)
Similarly, using $E[X(X-1)] = n(n-1)p^2$, then $V(X) = E[X^2] - (E[X])^2 = np(1-p)$

Question 5

(a) Credit Card Purchases

$X \sim \text{Binomial}(n = 10, p = 0.75)$

(i) $E(X) = np = 10 \times 0.75 = 7.5$
(ii) $V(X) = np(1-p) = 10 \times 0.75 \times 0.25 = 1.875$
(iii) $\sigma_X = \sqrt{V(X)} \approx 1.369$
(iv) $P(\mu - \sigma \leq X \leq \mu + \sigma) = P(6.131 \leq X \leq 8.869)$
Since X is discrete: $P(7 \leq X \leq 8)$
$P(X = 7) = \binom{10}{7}(0.75)^7(0.25)^3 \approx 0.2503$
$P(X = 8) = \binom{10}{8}(0.75)^8(0.25)^2 \approx 0.2816$
Total ≈ 0.5319 or 53.19%

(b) Book Checkout Time

(i) $P(0.5 \leq X \leq 1) = F(1) - F(0.5) = \frac{1^2}{4} - \frac{0.5^2}{4} = 0.25 - 0.0625 = 0.1875$
(ii) Median f satisfies F(f) = 0.5
$\frac{f^2}{4} = 0.5 \Rightarrow f^2 = 2 \Rightarrow f = \sqrt{2} \approx 1.414$ hours
(iii) Density function f(x)
$f(x) = F'(x) = \frac{x}{2}$ for $0 \leq x \leq 2$, 0 otherwise

(c) Water Dispensing

$X \sim N(64, 0.78^2)$
Find c such that $P(X > c) = 0.005$
$P(X \leq c) = 0.995 \Rightarrow Z \approx 2.5758$
$c = \mu + Z\sigma = 64 + 2.5758 \times 0.78 \approx 66.01$ oz

Question 6

(a) Asparagus Quality

(i) Sample Correlation Coefficient

Calculations: $\sum x = 707$, $\sum y = 23.59$, $\sum xy = 1759.11$
$\sum x^2 = 54961$, $\sum y^2 = 56.3135$
$n = 10$

$$
\begin{aligned}
r &= \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{[n\sum x^2 - (\sum x)^2][n\sum y^2 - (\sum y)^2]}} \\
&\approx \frac{10 \times 1759.11 - 707 \times 23.59}{\sqrt{(10 \times 54961 - 707^2)(10 \times 56.3135 - 23.59^2)}} \\
&\approx \frac{17591.1 - 16678.13}{\sqrt{(549610 - 499849)(563.135 - 556.4881)}} \\
&\approx \frac{902.97}{\sqrt{49761 \times 6.6469}} \approx \frac{902.97}{575.11} \approx 0.785
\end{aligned}
$$

Moderately strong positive linear relationship

(ii) Shear Force in Pounds

1 kg $\approx$ 2.20462 lbs
Correlation coefficient r is invariant to linear transformations, so it would remain 0.785

(b) Cheese Elongation

(i) Regression Line

Calculations: $n = 7$, $\sum x = 497$, $\sum y = 1219$, $\sum xy = 87129$, $\sum x^2 = 35767$

$$
\begin{aligned}
b &= \frac{n\sum xy - (\sum x)(\sum y)}{n\sum x^2 - (\sum x)^2} = \frac{7 \times 87129 - 497 \times 1219}{7 \times 35767 - 497^2} \\
&= \frac{609903 - 605843}{250369 - 247009} = \frac{4060}{3360} \approx 1.208
\end{aligned}
$$

$a = \bar{y} - b\bar{x} = \frac{1219}{7} - 1.208 \times \frac{497}{7} \approx 174.14 - 85.83 \approx 88.31$
Equation: $\hat{y} = 88.31 + 1.208x$

(ii) Estimate at x = 70

$\hat{y} = 88.31 + 1.208 \times 70 \approx 88.31 + 84.56 \approx 172.87\%$

(c) Piston Ring Diameter

$\mu = 12$ cm, $\sigma = 0.04$ cm, $n = 16$

(i) Center of Sampling Distribution

$E[\bar{X}] = \mu = 12$ cm

(ii) Standard Deviation of $\bar{X}$

$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} = \frac{0.04}{4} = 0.01$ cm
(iii) $P(\bar{X} > 12.01)$
$Z = \frac{12.01-12}{0.01} = 1$
$P(Z > 1) = 1 - \Phi(1) \approx 1 - 0.8413 = 0.1587$ or 15.87%