Econ 102_A01-A04 (Analysis of Economic Data): Cameron Fall 2021
Solutions to First Midterm Exam

Version A

1.(a) `Histogram, discrete` is best (as variable takes only six different values).

(b) For a Stata data set (more precisely Stata formatted dataset, usually with extension `.dta`).

(c) Take the natural logarithm (or `ln(x)`).

(d) The variable y is a standardized z-score with sample mean 0 & sample standard deviation 1.

(e) $\sum_{i=1}^{3}(1+\frac{6}{i}) = (1+\frac{6}{1})+(1+\frac{6}{2})+(1+\frac{6}{3}) = 7+4+3 = 14.$

(f) $\bar{x} = (3+7+8)/3 = 6.$ $s^2 = \frac{1}{2}\{(3-6)^2+(7-6)^2+(8-6)^2\} = \frac{1}{2}\{9+1+4\} = 14/2 = 7.$

2.(a) Lower quartile is 252.725.

(b) The smoothed histogram (not printed in these solutions) will be very right skewed as `skewness=4.278` is positive and much greater than zero and/or `mean=1656` is much greater than `median=803.`

(c) A 90 percent confidence interval is $1656.01 \pm t_{1064,.05} \times 2488.079/\sqrt{1065} =$
$= 1656.01 \pm 1.646 \times 76.241 = 1656.01 \pm 125.49 = (1530.5, 1781.5).$

(d) `mean spending`

(e) $H_0 : \mu = 1500$ against $H_0 : \mu \neq 1500$ at level 0.05.
$t = (1656.01 - 1500)/(2488.079/\sqrt{1065}) = 156.01/76.241 = 2.046.$
Since $|t| = 2.046 > t_{1064,.025} = 1.962$ reject H_0.
We reject $H_0 : \mu = 1500$ at significance level .05.

3.(a) One point each. (1) Common mean $\mathrm{E}[X_i] = \mu$; (2) Common variance $\mathrm{Var}[X_i] = \sigma^2$;
(3) Independence of the X_i.

(b) $\mu = \mathrm{E}[X] = 0.5 \times 10 + 0.2 \times 20 + 0.3 \times 30 = 5 + 4 + 9 = 18.$
$\sigma^2 = \mathrm{E}[(X-\mu)^2] = 0.5 \times (10-18)^2 + 0.2 \times (20-18)^2 + 0.3 \times (30-18)^2$
$= 0.5 \times 64 + 0.2 \times 4 + 0.3 \times 144 = 32 + 0.8 + 43.2 = 76.$

(c) The mean of variable `u` is 0.5 since it is uniformly distributed on the interval (0,1).

(d) Not surprising. For a 95% confidence interval we expect roughly 95% of 1000 or 950 to include μ so 50 do not include μ. 60 is quite close.

(e) $\bar{X}$ has mean 200, variance $100/25 = 4$ and standard deviation $\sqrt{4} = 2$.

Multiple Choice for Version A

Question	**1.**	**2.**	**3.**	**4.**	**5.**
Answer Version A	*c*	*b*	*d*	*a*	*b*

3. is $100 \pm 2 \times \sqrt{25}$.

The course grade will be based on a curve from the combined scores of midterm 1 (20%), midterm 2 (20%), final (40%) and assignments (20%). **The curve for this exam is only a guide**. Curve below has average GPA 2.72 for this exam.

Scores out of	35
75*th* percentile	30 (86%)
Median	26 (74%)
25*th* percentile	22 (63%)

A+	34 and above	C+	23 and above
A	30 and above	C	22and above
A-	28 and above	C-	20 and above
B+	27 and above	D+	19 and above
B	26 and above	D	18 and above
B-	25 and above	D-	17 and above

Version B of 102A

1.(a) $\sum_{i=1}^{3}(1+6i) = (1+6\times 1)+(1+6\times 2)+(1+6\times 3) = 7+13+19 = 39.$

(b) The variable y is a standardized z-score with sample mean 0 & sample standard deviation 1.

(c) $\bar{x} = (3+7+2)/3 = 4.$ $s^2 = \frac{1}{2}\{(3-4)^2+(7-4)^2+(2-4)^2\} = \frac{1}{2}\{1+9+4\} = 14/2 = 7.$

(d) For a Stata data set (more precisely Stata formatted dataset, usually with extension `.dta`).

(e) Take the natural logarithm (or `ln(x)`).

(f) `Histogram, discrete` is best (as variable takes only six different values).

2.(a) Upper quartile is 1995.528.

(b) The smoothed histogram (not printed in these solutions) will be very right skewed as `skewness=4.328` is positive and much greater than zero and/or `mean=1692` is much greater than `median=807.`

(c) A 99 percent confidence interval is $1692.75 \pm t_{920,.005} \times 2566.21/\sqrt{921} =$
$= 1692.75 \pm 2.581 \times 84.559 = 1692.75 \pm 218.25 = (1474.5, 1911.0).$

(d) `mean spending`

(e) $H_0 : \mu = 1600$ against $H_0 : \mu \neq 1600$ at level 0.05.
$t = (1692.75 - 1600)/(2566.21/\sqrt{921}) = 92.75/84.559 = 1.097.$
Since $|t| = 1.097 < t_{920,.025} = 1.963$ do not reject H_0.
We do not reject $H_0 : \mu = 1600$ at significance level .05.

3.(a) $\mu = \mathrm{E}[X] = 0.3\times 10 + 0.3\times 20 + 0.4\times 30 = 3+6+12 = 21.$
$\sigma^2 = \mathrm{E}[(X-\mu)^2] = 0.3\times(10-21)^2 + 0.3\times(20-21)^2 + 0.4\times(30-21)^2$
$= 0.3\times 121 + 0.3\times 1 + 0.4\times 81 = 36.3 + 0.3 + 32.4 = 69.$

(b) The mean of variable `x` is $\frac{1}{6} = 0.1667$ since $x = 1$ with probability $\frac{1}{6}$ and $x = 0$ with probability $\frac{5}{6}$.

(c) $\bar{X}$ has mean 200, variance $900/100 = 9$ and standard deviation $\sqrt{9} = 3$.

(d) Surprising. For a 95% confidence interval we expect roughly 95% of 1000 or 950 to include μ. 40 is very different from 950.

(e) One point each. (1) Common mean $\mathrm{E}[X_i] = \mu$; (2) Common variance $\mathrm{Var}[X_i] = \sigma^2$;
(3) Independence of the X_i.

Multiple Choice for Version B

Question	**1.**	**2.**	**3.**	**4.**	**5.**
Answer Version B	*b*	*b*	*c*	*d*	*b*

4. is $400 \pm 2 \times \sqrt{100}$.

The course grade will be based on a curve from the combined scores of midterm 1 (20%), midterm 2 (20%), final (40%) and assignments (20%). **The curve for this exam is only a guide**. Curve below has average GPA 2.72 for this exam.

Scores out of	35
75*th* percentile	30 (86%)
Median	26 (74%)
25*th* percentile	22 (63%)

A+	34 and above	C+	23 and above
A	30 and above	C	22and above
A-	28 and above	C-	20 and above
B+	27 and above	D+	19 and above
B	26 and above	D	18 and above
B-	25 and above	D-	17 and above