Econ 102 A01-A04 (Analysis of Economic Data): Cameron Fall 2022
Solutions to First Midterm Exam

Version A

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

(b) No. (The command `use` can only read 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) $\sum_{i=1}^{3}(5+2i^2) = (5+2\times 1^2)+(5+2\times 2^2)+(5+2\times 3^2) = 7+13+23 = 43$.

(f) $\bar{x} = (1+2+3+6)/4 = 3$.
$s^2 = \frac{1}{3}\{(1-3)^2+(2-3)^2+(3-3)^2+(6-3)^2\} = \frac{1}{3}\{4+1+0+9\} = 14/3 = 4.667$.

2.(a) Lower quartile is 30.

(b) No. There is very mild skewness. `skewness=0.389` is not too different from zero and `mean=39.6` is close to `median=37.`

(c) Box plot.

(d) A 90 percent confidence interval is $39.632 \pm t_{1193,.05} \times 10.586/\sqrt{1194} =$
$= 39.632 \pm 1.646 \times 0.3064 = 39.632 \pm 0.5043 = (39.13, 40.14)$.

(e) `mean age`

(f) $H_0 : \mu = 39$ against $H_0 : \mu \neq 39$ at level 0.05.
$t = (39.632-39)/(10.586/\sqrt{1194}) = 0.632/0.3064 = 2.063$.
Since $|t| = 2.063 > t_{1193,.025} = 1.962$ reject H_0.
We reject $H_0 : \mu = 39$ 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) $\mu = \mathrm{E}[X] = 0.1\times 2 + 0.6\times 5 + 0.3\times 6 = 0.2+3.0+1.8 = 5$.
$\sigma^2 = \mathrm{E}[(X-\mu)^2] = 0.1\times(2-5)^2 + 0.3\times(5-5)^2 + 0.3\times(6-5)^2$
$= 0.1\times 9 + 0.3\times 0 + 0.3\times 1 = 0.9+0+0.3 = 1.2$.

(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*	*c*	*a*	*b*

3. is $100 \pm 1 \times \sqrt{25}$.

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 for this exam has average GPA 2.79.

Scores out of	35
75th percentile	31.5 (90%)
Median	30 (86%)
25th percentile	26.5 (76%)

A+	34 and above	C+	26.5 and above
A	32 and above	C	25.5 and above
A-	31and above	C-	24 and above
B+	30 and above	D+	22 and above
B	29 and above	D	20 and above
B-	28 and above	D-	18 and above

Version B of 102A

1.(a) $\bar{x} = (1+5+6+8)/4 = 5$.
$s^2 = \frac{1}{3}\{(1-5)^2+(5-5)^2+(6-5)^2+(8-5)^2\} = \frac{1}{3}\{16+0+1+9\} = 26/3 = 8.667$.

(b) $\sum_{i=1}^{3}(2+\frac{6}{i}) = (2+\frac{6}{1})+(2+\frac{6}{2})+(2+\frac{6}{3}) = 8+5+4 = 17$

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

(d) No. (The command `use` can only read a Stata data set (more precisely Stata formatted dataset, usually with extension `.dta`).

(e) `Histogram, discrete` is best (as variable takes only four different values).

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

2.(a) Upper quartile is 49.

(b) No. There is very mild skewness. `skewness=0.355` is not too different from zero and `mean=39.8` is close to `median=39.`

(c) Box plot.

(d) A 99 percent confidence interval is $39.776 \pm t_{919,.005} \times 10.675/\sqrt{920} =$
$= 39.776 \pm 2.581 \times 0.3519 = 39.776 \pm 0.908 = (38.87, 40.68)$.

(e) `mean age`

(f) $H_0: \mu = 40$ against $H_0: \mu \neq 40$ at level 0.05.
$t = (39.776-40)/(10.675/\sqrt{920}) = -0.224/0.3519 = -0.637$.
Since $|t| = 0.637 < t_{919,.025} = 1.963$ do not reject H_0.
We do not reject $H_0: \mu = 40$ at significance level .05.

3.(a) $\mu = \mathrm{E}[X] = 0.3\times 2 + 0.5\times 4 + 0.2\times 7 = 0.6+2+1.4 = 4$.
$\sigma^2 = \mathrm{E}[(X-\mu)^2] = 0.3\times(2-4)^2 + 0.5\times(4-4)^2 + 0.2\times(7-4)^2$
$= 0.3\times 4 + 0.5\times 0 + 0.2\times 9 = 1.2+0+1.8 = 3$.

(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 300, variance $400/16 = 25$ and standard deviation $\sqrt{25} = 5$

(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	*d*	*a*	*c*	*a*	*b*

1. is $400 \pm 1 \times \sqrt{100}$.

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 for this exam has average GPA 3.79.

Scores out of	35
75*th* percentile	31.5 (90%)
Median	30 (86%)
25*th* percentile	26.5 (76%)

A+	34 and above	C+	26.5 and above
A	32 and above	C	25.5 and above
A-	31and above	C-	24 and above
B+	30 and above	D+	22 and above
B	29 and above	D	20 and above
B-	28 and above	D-	18 and above