Problem Set 1:
Probability and Statistics

Marek Chadim
Econometrics I

January 26, 2024

1.

$$\begin{aligned}
\text{Var}(X+Y) &= E\left[(X+Y)^2\right] - (E[X+Y])^2 \\
&= E\left[X^2 + 2XY + Y^2\right] - (E[X] + E[Y])^2 \\
&= \text{Var}(X) + \text{Var}(Y) + 2\left(E[XY] - E[X]E[Y]\right) \\
&= \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X,Y)
\end{aligned}$$

2. (a) i. Expectation of u: $E[u] = 0$ for a uniform distribution $U[-1,1]$.

ii. Expectation of v ($E[v] = E[u^2]$):

$$\begin{aligned}
E[u^2] &= \int_{-1}^{1} u^2 \cdot \frac{1}{2}\, du = \frac{1}{2}\left[\frac{u^3}{3}\right]_{-1}^{1} \\
&= \frac{1}{2}\left(\frac{1}{3} + \frac{1}{3}\right) = \frac{1}{3}
\end{aligned}$$

iii. Expectation of uv ($E[uv] = E[u^3]$):

$$E[u^3] = \int_{-1}^{1} u^3 \cdot \frac{1}{2}\, du = \frac{1}{2}\left[\text{Since } u^3 \text{ is an odd function}\right] = 0$$

iv. Covariance:

$$\text{Cov}(u,v) = E[uv] - E[u]E[v] = 0 - 0 \cdot \frac{1}{3} = 0$$

(b) Define $A = \{u : u > \frac{1}{2}\}$ and $B = \{v : v \leq \frac{1}{4}\}$. The event B is equivalent to $u \in [-\frac{1}{2}, \frac{1}{2}]$. Then $P(A) = \frac{1}{4}$, and $P(B) = \frac{1}{2}$. However, $A \cap B$ is the event where $u > \frac{1}{2}$ and $u \in [-\frac{1}{2}, \frac{1}{2}]$ As such, $P(A \cap B) = 0 \neq P(A)P(B) = \frac{1}{4} \times \frac{1}{2} = \frac{1}{8}$.

(c) For $v_0 = \frac{1}{4}$, u can be either $\frac{1}{2}$ or $-\frac{1}{2}$. The probability of u being less than any value other than $\pm\frac{1}{2}$ given $v = \frac{1}{4}$ is either 0 or 1, differing from $F_u(u) = \frac{u+1}{2}, \quad -1 \leq u \leq 1$.

3. (a) Since expectation is a linear operator, $E(\bar{y}) = E\left(\frac{1}{n}\sum_{i=1}^{n} y_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(y_i) = \frac{1}{n} \cdot n\mu = \mu$ and $\text{Bias}(\bar{y}) = E(\bar{y}) - \mu = \mu - \mu = 0$.

(b) Since y_1 is an observation from the distribution with mean μ, $E(y_1) = \mu$. Therefore, $\text{Bias}(y_1) = E(y_1) - \mu = \mu - \mu = 0$

(c) $\text{MSE}(\bar{y}) = \text{Var}(\bar{y}) + [\text{Bias}(\bar{y})]^2 = \frac{1}{n^2}\sum_{i=1}^{n} \text{Var}\,(y_i) = \frac{\sigma^2}{n}$
$\text{MSE}(y_1) = \text{Var}(y_1) + [\text{Bias}(y_1)]^2 = \sigma^2$

(d) Both estimators are unbiased, yet $\bar{y}$ is preferred over y_1 as it is more efficiency, with a lower MSE for sample sizes greater than 1.

(e) $E[s^2] = \frac{1}{n}\sum_{i=1}^{n} E[(y_i - \mu)^2] + 2E[(y_i - \mu)(\mu - \bar{y})] + E[(\mu - \bar{y})^2]$
$E[(y_i - \mu)^2] = \sigma^2$, $E[(y_i - \mu)(\mu - \bar{y})] = 0$, $E[(\mu - \bar{y})^2] = \text{Var}(\bar{y}) = \frac{\sigma^2}{n}$

$$\text{Bias}(s^2) = E[s^2] - \sigma^2 = \frac{-\sigma^2}{n}$$

(f) $\frac{1}{2}E[y_1^2 - 2y_1y_2 + y_2^2] = \frac{1}{2}(E[y_1^2] - 2E[y_1y_2] + E[y_2^2]) = [\sigma^2 + \mu^2 - \mu^2]$

$$\text{Bias}(\hat{\sigma}^2) = E[\hat{\sigma}^2] - \sigma^2 = 0$$

(g) $E(s^2 - \sigma^2)^2 = \text{Var}(s^2) + \text{Bias}^2 = \frac{2(n-1)\sigma^4}{n^2} + \left(\frac{-\sigma^2}{n}\right)^2 = \frac{2n-1}{n^2}\sigma^4$.
$E(\hat{\sigma}^2 - \sigma^2)^2 = \text{Var}(\hat{\sigma}^2) = \frac{1}{16}E((y_1 - y_2)^8) - [E(\hat{\sigma}^2)]^2 = 8\sigma^8 - \sigma^4$

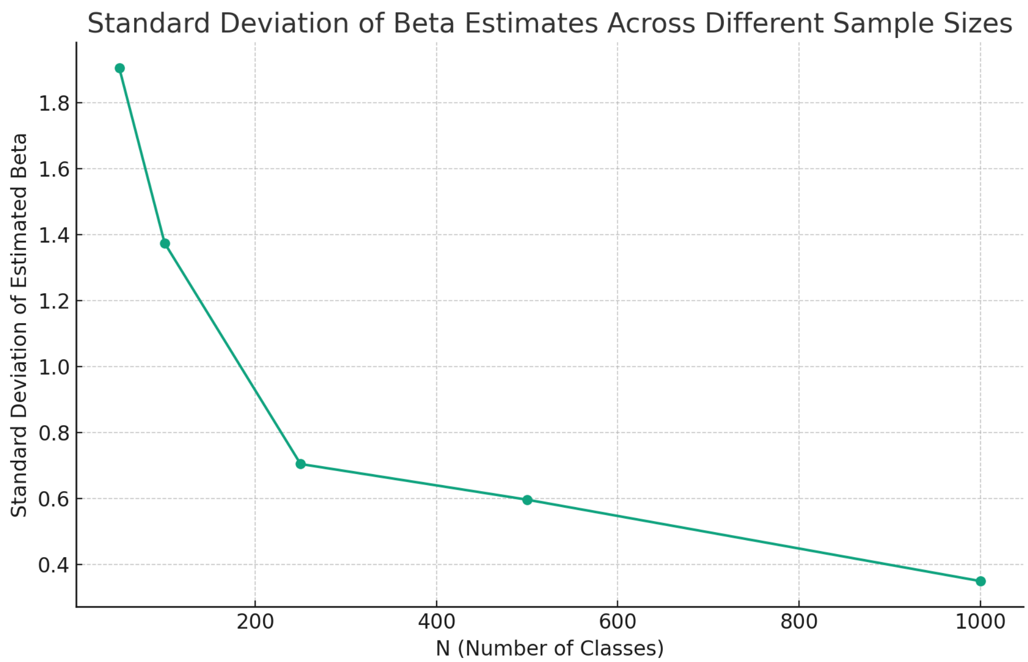
(h) s^2 is preferred (biased, yet consistent with MSE decreasing in n)

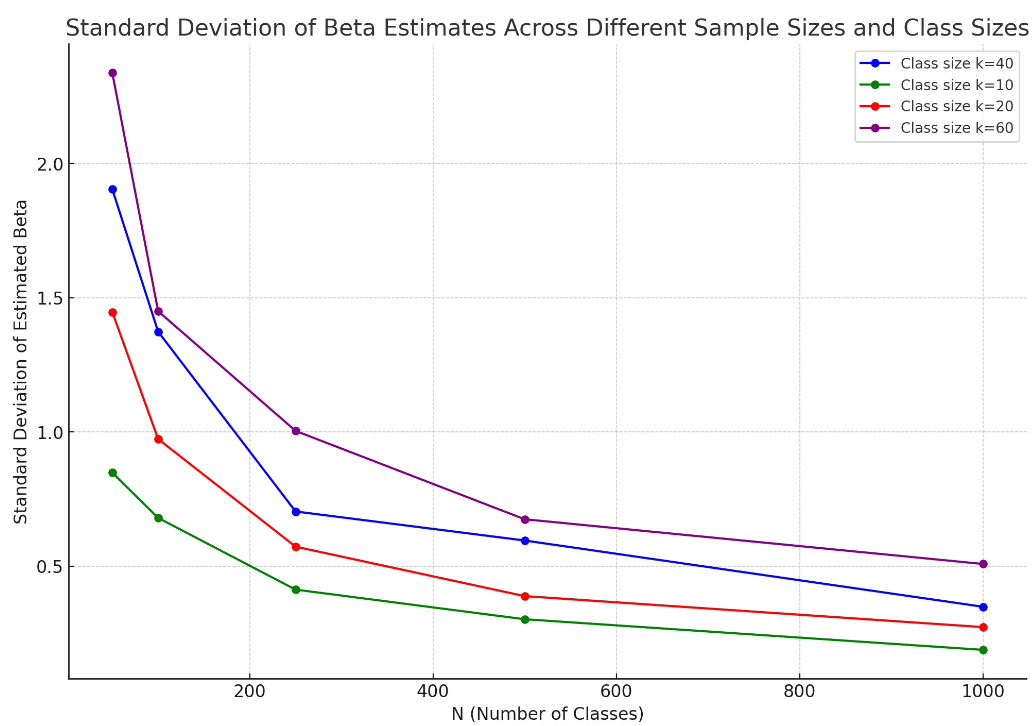
4. (a) $\sigma_{s_c} = \sqrt{\frac{0.556 \times (1 - 0.556)}{102}} \approx 0.0492 <$ reported 0.118

(b) i. below

ii. below

iii. Increasing the number of classes (larger sample) and decreasing the class size (more variation) led to lower standard deviations, in line with the rule of thumb.


Standard Deviation of Beta Estimates Across Different Sample Sizes
Standard Deviation of Estimated Beta
N (Number of Classes)
0.4
0.6
0.8
1.0
1.2
1.4
1.6
1.8
200
400
600
800
1000


Standard Deviation of Beta Estimates Across Different Sample Sizes and Class Sizes
Class size k=40
Class size k=10
Class size k=20
Class size k=60
Standard Deviation of Estimated Beta
N (Number of Classes)
0.5
1.0
1.5
2.0
200
400
600
800
1000

iv. Current Sample: 200 classes, 30 students/class. Urban Option: +100 classes, 40 students/class. Rural Option: +50 classes, 15 students/class. Simulation Results:
$\sigma_{\hat{\beta},\text{Urban}} \approx 1.053$, $\sigma_{\hat{\beta},\text{Rural}} \approx 0.975$
Rural expansion is likely to provide more precise estimates.

(c) i. $F(k; n, \frac{1}{2}) = \sum_{i=0}^{k} \binom{n}{i} \left(\frac{1}{2}\right)^i \left(\frac{1}{2}\right)^{n-i}$, gives the probability of k or fewer female professors after taking n courses.

ii. $P(X \leq 2) + P(X \geq 7) = 2\sum_{i=0}^{2} \binom{10}{i} \left(\frac{1}{2}\right)^i \left(\frac{1}{2}\right)^{10-i} = 0.109375$

iii. $P(X \leq 4) + P(X \geq 16) = 2\sum_{i=0}^{4} \binom{20}{i} \left(\frac{1}{2}\right)^i \left(\frac{1}{2}\right)^{20-i} \approx 0.0047$

iv. $P(X \leq 14)+P(X \geq 46) = 2\sum_{i=0}^{14} \binom{60}{i} \left(\frac{1}{2}\right)^i \left(\frac{1}{2}\right)^{60-i} \approx 8{*}10^{-10}$

5. (a) $E(x) = \lambda \int_0^\infty xe^{-\lambda x}\,dx = \lambda \left(x\left(-\frac{1}{\lambda}e^{-\lambda x}\right)\Big|_0^\infty - \int_0^\infty -\frac{1}{\lambda}e^{-\lambda x}\,dx \right) = \lambda \left(\frac{1}{\lambda}\left(-\frac{1}{\lambda}e^{-\lambda x}\right)\Big|_0^\infty \right) = \frac{1}{\lambda}$

(b) $F(x) = \int_0^x \lambda e^{-\lambda t}dt = \left[-e^{-\lambda t}\right]_0^x = -e^{-\lambda x} - (-e^{-\lambda \cdot 0}) = 1 - e^{-\lambda x}$

(c) The MGF of $y(n)$ is

$$M_{y(n)}(t) = E(e^{ty(n)}) = E(e^{t(x-\bar{x}_n)})$$

The MGF of $\tilde{x}$ is

$$M_{\tilde{x}}(t) = E(e^{t\tilde{x}}) = E(e^{t(x-E(x))})$$

By the law of large numbers

$$\lim_{n\to\infty} M_{y(n)}(t) = E(e^{t(x-E(x))}) = M_{\tilde{x}}(t)$$
$$\iff \lim_{n\to\infty} |F_{y(n)}(y) - F_{\tilde{x}}(y)| = 0$$

(d) Define $\hat{p}_n = \frac{1}{n}\sum_{i=1}^{n} 1\{x_i \leq 1\}$. Let $p = Pr(x \leq 1)$, $\epsilon > 0$. By Chebyshev's inequality with $k = \frac{\epsilon n}{\sqrt{p(1-p)}}$

$$P\left(|\hat{p}_n - p| \geq \epsilon\right) \leq \frac{p(1-p)}{n\epsilon^2}$$

Thus

$$\lim_{n\to\infty} P\left(|\hat{p}_n - p| \geq \epsilon\right) = 0,\ \forall \epsilon > 0.$$