STAT390: Homework 2

Ken

2023-10-11

Homework 3

①

a) pmf of X_i

$P(X_i = 3) = \frac{1}{3}$

$P(X_i = 9) = \frac{1}{3}$

$P(X_i = 0) = \frac{1}{3}$

b) pmf of $\bar{X}$

$P(\bar{X} = 0) = \{000\} = \frac{1}{27}$

$P(\bar{X} = 1) = \{300, 030, 003\} = \frac{3}{27}$

$P(\bar{X} = 2) = \{330, 303, 003\} = \frac{3}{27}$

$P(\bar{X} = 3) = \{900, 090, 009, 333\} = \frac{4}{27}$

$P(\bar{X} = 4) = \{390, 309, 039, 930, 903, 093\} = \frac{6}{27}$

$P(\bar{X} = 5) = \{339, 393, 933\} = \frac{3}{27}$

$P(\bar{X} = 6) = \{990, 909, 099\} = \frac{3}{27}$

$P(\bar{X} = 7) = \{993, 939, 399\} = \frac{3}{27}$

$P(\bar{X} = 8) = \frac{0}{27}$

$p(\bar{x}=9) = \{999\} = \frac{1}{27}$

c) cdf of $\bar{x}$

$$F(a) = \begin{cases} 0 & a < 0 \\ 1/27 & 0 \le a < 1 \\ 4/27 & 1 \le a < 2 \\ 7/27 & 2 \le a < 3 \\ 11/27 & 3 \le a < 4 \\ 17/27 & 4 \le a < 5 \\ 20/27 & 5 \le a < 6 \\ 23/27 & 6 \le a < 7 \\ 26/27 & 7 \le a < 8 \\ 26/27 & 8 \le a < 9 \\ 1 & a \ge 9 \end{cases}$$

d) $E[x_1] = \sum x_1 \frac{1}{3}$

$3\frac{1}{3} + 9\frac{1}{3} + 0\frac{1}{3} = 4$

$\boxed{E[x_1] = 4}$

$\sqrt{E[x_1^2] - (E[x_1])^2}$

$\frac{9}{3} + \frac{81}{3} + 0 = 30$

$\boxed{SD(x_1) = \sqrt{30 - 16} = \sqrt{14}}$

$E[\bar{x}] = \sum_{i=9} \bar{x}_i \, p(\bar{x}_i)$

$0 \cdot \frac{1}{27} + 1 \cdot \frac{3}{27} + 2 \cdot \frac{3}{27} + 3 \cdot \frac{4}{27} +$
$4 \cdot \frac{6}{27} + 5 \cdot \frac{3}{27} + 6 \cdot \frac{3}{27} + 7 \cdot \frac{3}{27} +$
$8 \cdot \frac{0}{27} + 9 \cdot \frac{1}{27} = 4$

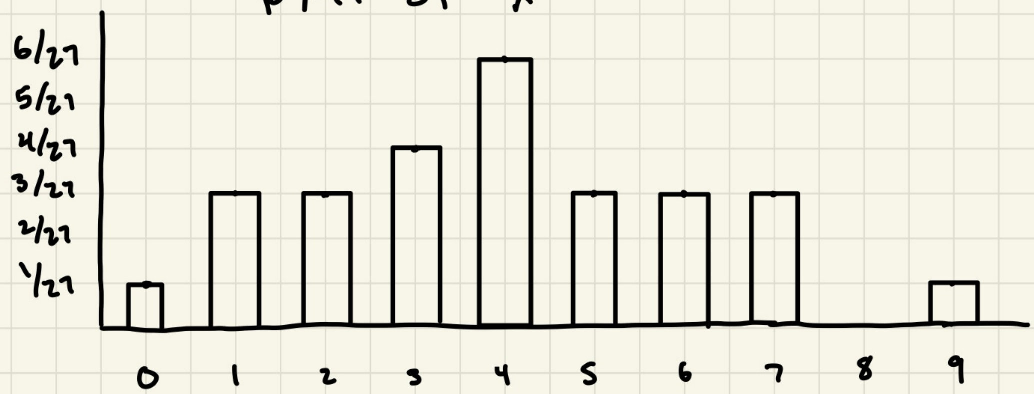
$\boxed{E[\bar{x}] = 4}$

$\sqrt{E[\bar{x}^2] - E[\bar{x}]^2}$

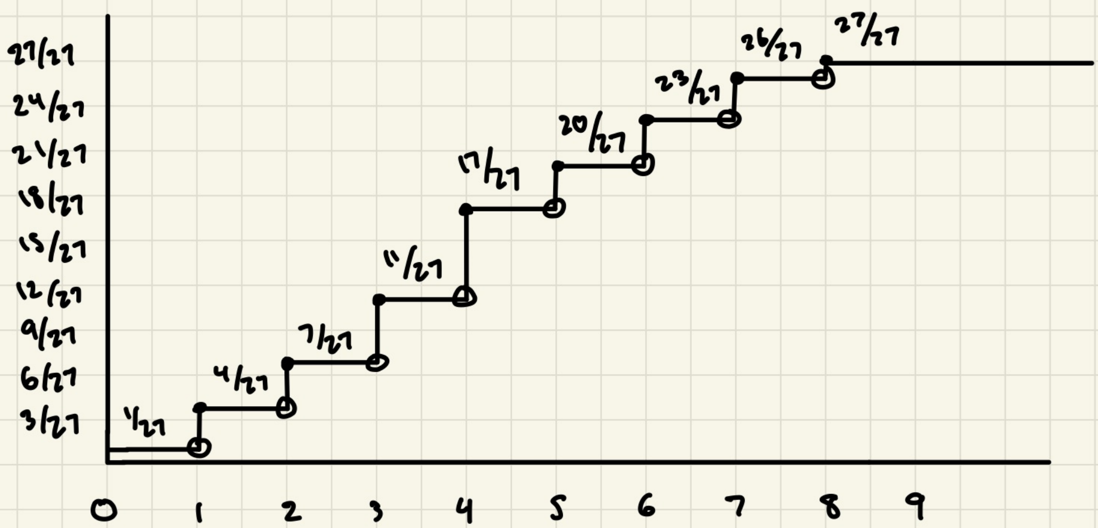
$\boxed{SD(\bar{x}) = \sqrt{20\frac{2}{3} - 16} = \sqrt{\frac{14}{3}}}$

Both have the same expected value but different standard deviation.

e) PMF of $\bar{X}$



CDF of $\bar{X}$



2

a) $\boxed{\text{Binomial Distribution} \quad \text{Bin}(2000, .0015)}$

b) $\boxed{E[X] = 3}$

$SD(X) = \sqrt{np(1-p)} = \boxed{\sqrt{3(1-.0015)}}$

c) $P(X=0)$, $P(X=1)$, $P(X>2)$

$= 1 - P(X \leq 2)$

$= 1 - P(X=2) - P(X=1) - P(X=0)$

$P(X=0) = \binom{2000}{0} .0015^0 (1-.0015)^{2000-0} = \boxed{.0497}$

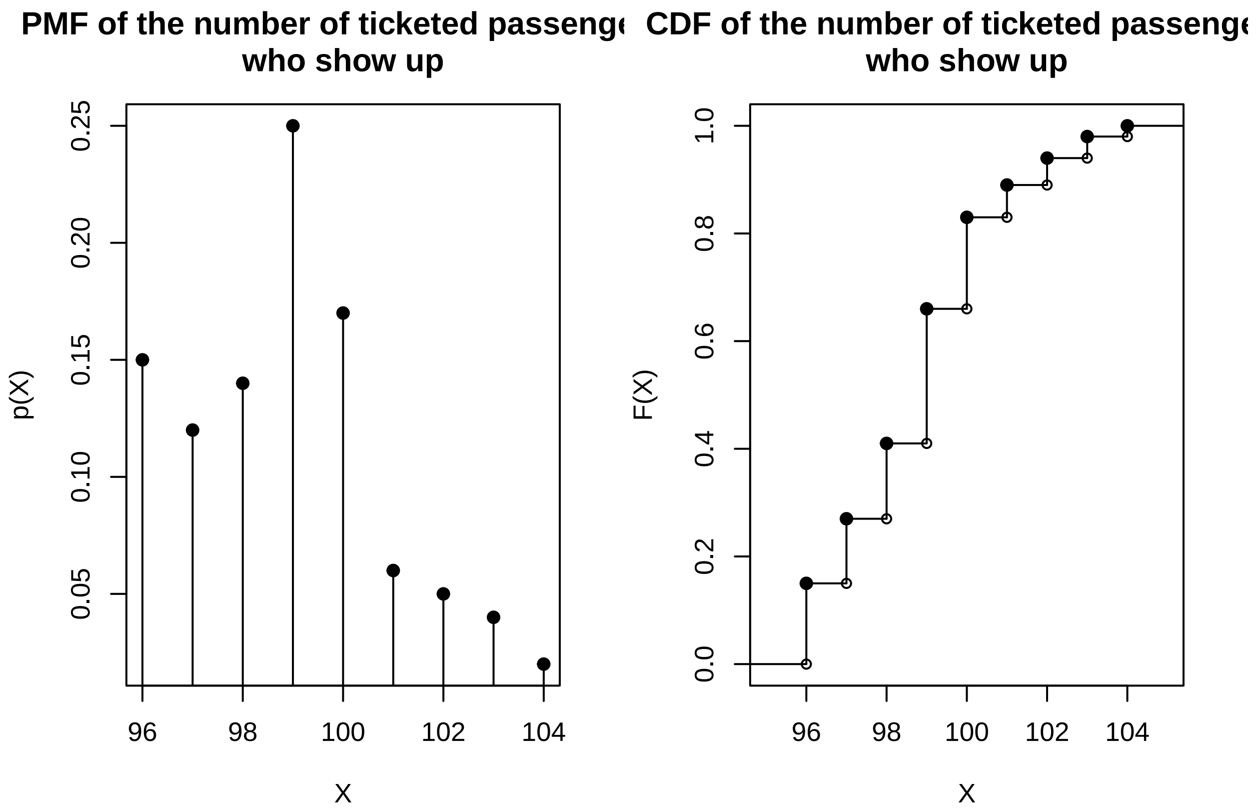
$P(X=1) = \binom{2000}{1} .0015^1 (1-.0015)^{2000-1} = \boxed{.1492}$

$P(X=2) = \binom{2000}{2} .0015^2 (1-.0015)^{2000-2} = .2241$

$P(X>2) = 1 - .0497 - .1492 - .2241 = \boxed{.577}$

Question 3

```
x <- c(96, 97, 98, 99, 100, 101, 102, 103, 104)
pmf <- c(.15, .12, .14, .25, .17, .06, .05, .04, .02)
cdf <-  stepfun(x, c(0, .15, .27, .41, .66, .83, .89, .94, .98, 1), right=TRUE)
par(mfrow=c(1,2), cex=.8)
plot(x, pmf, type = "h", main = "PMF of the number of ticketed passengers\nwho show up", xlab = "X", yla
points(x, pmf, pch = 19)
plot(cdf, main = "CDF of the number of ticketed passengers\nwho show up", xlab = "X", ylab = "F(X)")
points(x, c(.15, .27, .41, .66, .83, .89, .94, .98, 1), pch = 19)
```



③ b) $p(X \le 100) = \boxed{.83}$

$p(98 < X \le 102) = .94 - .41 = \boxed{.53}$

c) $|X - 100| = 4, 3, 2, 1, 0, 1, 2, 3, 4$

X	96	97	98	99	100	101	102	103	104
Y	4	3	2	1	0	1	2	3	4

Y	0	1	2	3	4
f(Y)	.17	.31	.19	.16	.17

$E[Y] = .31 + .19 \cdot 2 + .16 \cdot 3 + .17 \cdot 4 = \boxed{1.85}$

d)

X	96	97	98	99	100	101	102	103	104
Z	800	800	800	800	800	500	200	-100	-400

Z	-400	-100	200	500	800
Z(x)	.02	.04	.05	.06	.83

$E[Z] = -400 \cdot .02 - 100(.04) + 200(.05) + 500(.06) + 800(.83)$

$\boxed{E[Z] = 692}$

$\sqrt{E[Z^2] - E[Z]^2} = SD(Z)$

$\sqrt{551800 - 478864} = \boxed{270.07 = SD(Z)}$

Question 4

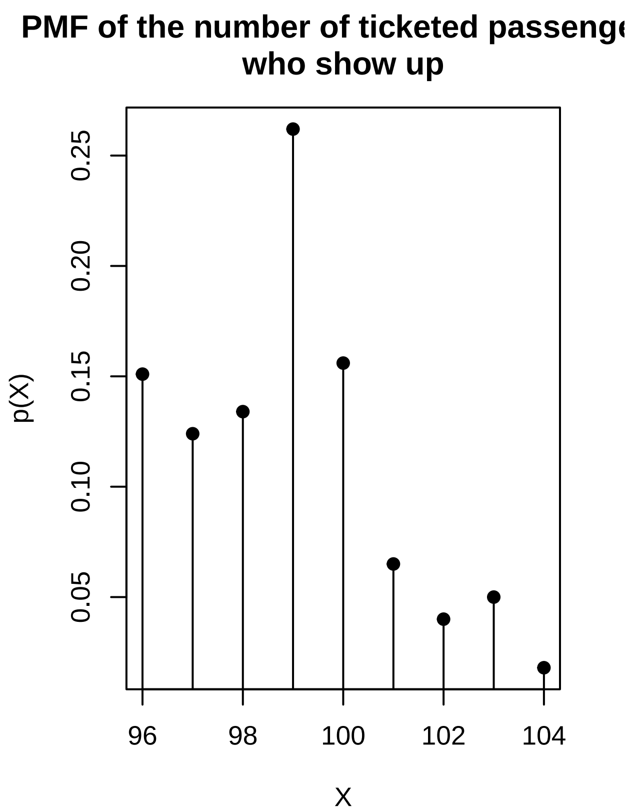
```
p <- c(.15, .12, .14, .25, .17, .06, .05, .04, .02)
X <- sample(x, 1000, replace=T, prob=p)
Z <- 800-300*(X-100)*(X > 100)
```

First 10 simulations: [200, 800, 800, 800, 800, 800, 800, 500, 800, 500]

```
counts <- c(length(X[X == 96]) / 1000, length(X[X == 97]) / 1000, length(X[X == 98]) / 1000, length(X[X
par(mfrow=c(1,2), cex=.8)
plot(x, counts, type = "h", main = "PMF of the number of ticketed passengers\nwho show up", xlab = "X",
points(x, counts, pch = 19)

prob1 <- length(X[X <= 100]) / 1000
prob2 <- length(X[X > 98 & X <= 102]) / 1000

mean <- mean(Z)
sd <- sd(Z)
```



P(X <= 100) is .84, P(98 < X <= 102) is .537, the mean is 698, and the standard deviation is 263.11. The histograms look the almost exactly the same, the probabilities are almost exactly the same, and the mean and standard deviation are very close as well.