

# Homework 1

Yunlin Zhang UID: 17583629

## 1 Problem 1

$X \sim \mathcal{N}(5, 4) \Rightarrow \mu = 5, \sigma^2 = 4 \Rightarrow \sigma = 2$

All numbers based on  $\mathcal{N}(0, 1)$  lookup table

$$\begin{aligned} \text{i. } & P[X \leq 6] \\ &= P[X \leq \mu + 0.5\sigma] \\ &\approx 0.6915 \end{aligned}$$

$$\begin{aligned} \text{ii. } & P[X > 4] \\ &= P[X > \mu - 0.5\sigma] \\ &= P[X < \mu + 0.5\sigma] \\ &\approx 0.6915 \end{aligned}$$

$$\begin{aligned} \text{iii. } & P[|X - 5| > 1] \\ &= P[X - 5 > 1 \vee X - 5 < -1] \\ &= P[X > 6 \vee X < 4] \\ &= P[X > \mu + 0.5\sigma] + P[X < \mu - 0.5\sigma] \\ &= 2P[X > \mu + 0.5\sigma] \\ &= 2(1 - P[X < \mu + 0.5\sigma]) \\ &\approx 0.6171 \end{aligned}$$

## 2 Problem 2

Probability of beating the market each year is 0.5

$$\text{i. } P[\text{beat market all 10 years}] = p^n = \left(\frac{1}{2}\right)^{10} = \frac{1}{1024} \approx 0.0977\%$$

$$\begin{aligned} \text{ii. Consider } n = 4170, p = 0.0977\%, q = 1 - 0.0977\% = 99.9023\% \\ P[\text{at least 1 beats market all 10 years}] &= 1 - P[\text{no one beat market all 10 years}] \\ &= 1 - \binom{4170}{0} (0.0977\%)^0 (99.9023\%)^{4170} \end{aligned}$$

$$\approx 1 - 0.017 = 0.983 = 98.3\%$$

Out of 4170 funds it is extremely likely that at least one will beat the market all 10 years.

$$\text{iii. } P[\text{at least 5 funds beats market all 10 years}] = \sum_{i=5}^{4170} \binom{4170}{i} (0.0977\%)^i (99.9023\%)^{4170-i} \\ \approx 38.5\%$$

MATLAB code:

---

```
p = (1/2)^10; q = 1 - p; s = 0; N = 4170;
for i = 5:N
    s = s + binopdf(i,N,p);
end
```

---

### 3 Problem 3

$Y_i$  are i.i.d,  $E[Y_i] = \mu$ ,  $Var[Y_i] = \sigma^2$  and  $\bar{Y} = \frac{1}{4} \sum_{i=1}^4 Y_i$

i. Calculate  $E[\bar{Y}]$  and  $Var[\bar{Y}]$

$$E[\bar{Y}] = E[\frac{1}{4} \sum_{i=1}^4 Y_i] = \frac{1}{4} \sum_{i=1}^4 E[Y_i] = \frac{1}{4}(4\mu)$$

$$= \mu$$

$$Var[\bar{Y}] = Var[\frac{1}{4} \sum_{i=1}^4 Y_i] = (\frac{1}{4})^2 \sum_{i=1}^4 Var[Y_i]$$

$\because Y_i$  are i.i.d

$$= \frac{1}{16}(4\sigma^2) = \frac{1}{4}\sigma^2$$

ii.  $W = \frac{1}{8}Y_1 + \frac{1}{8}Y_2 + \frac{1}{4}Y_3 + \frac{1}{2}Y_4$  calculate  $E[W]$  and  $Var[W]$

$$E[W] = E[\frac{1}{8}Y_1 + \frac{1}{8}Y_2 + \frac{1}{4}Y_3 + \frac{1}{2}Y_4] = \frac{1}{8}E[Y_1] + \frac{1}{8}E[Y_2] + \frac{1}{4}E[Y_3] + \frac{1}{2}E[Y_4]$$

$$= \frac{1}{8}\mu + \frac{1}{8}\mu + \frac{1}{4}\mu + \frac{1}{2}\mu$$

$$= \mu \Rightarrow W \text{ is an unbiased estimator of } \mu$$

$$Var[W] = Var[\frac{1}{8}Y_1 + \frac{1}{8}Y_2 + \frac{1}{4}Y_3 + \frac{1}{2}Y_4]$$

$$= (\frac{1}{8})^2 Var[Y_1] + (\frac{1}{8})^2 Var[Y_2] + (\frac{1}{4})^2 Var[Y_3] + (\frac{1}{2})^2 Var[Y_4]$$

$\because Y_i$  are i.i.d

$$= (\frac{1}{8})^2 \sigma^2 + (\frac{1}{8})^2 \sigma^2 + (\frac{1}{4})^2 \sigma^2 + (\frac{1}{2})^2 \sigma^2$$

$$= \frac{11}{32}\sigma^2 > \frac{1}{4}\sigma^2$$

iii.  $\bar{Y}$  would be a better estimator of  $\mu$  since its variance is less than that of  $W$ , therefore it is a more efficient estimator

### 4 Problem 4

Consider  $Y_i$ ,  $1 \leq i \leq n$ , and  $E[Y_i] = \mu$ ,  $Var[Y_i] = \sigma^2$ , and  $Cov[Y_i, Y_j] = 0$  for  $i \neq j$

i. Define  $W_a = \sum_{i=1}^n a_i Y_i$   
 $\Rightarrow E[W] = E[\sum_{i=1}^n a_i Y_i] = \sum_{i=1}^n E[a_i Y_i] = \sum_{i=1}^n a_i E[Y_i]$   
 $= \sum_{i=1}^n a_i \mu = \mu \sum_{i=1}^n a_i$   
 $W$  is an unbiased estimator of  $\mu \iff E[W] = \mu \Rightarrow \mu \sum_{i=1}^n a_i = \mu$   
 $\Rightarrow \sum_{i=1}^n a_i = 1$  and  $\mu \neq 0$

ii. Find  $Var[W]$   
 $Var[W] = Var[\sum_{i=1}^n a_i Y_i] = \sum_{i=1}^n Var[a_i Y_i] \quad \because Cov[Y_i, Y_j] = 0 \text{ for } i \neq j$   
 $= \sum_{i=1}^n a_i^2 Var[Y_i] = \sum_{i=1}^n a_i^2 \sigma^2$   
 $= \sigma^2 \sum_{i=1}^n a_i^2$

iii. Give  $\frac{1}{n}(\sum_i a_i)^2 \leq \sum_i a_i^2$  show  $\forall a \text{ s.t. } E[W] = \mu, Var[W_a] \geq Var[\bar{Y}]$   
 $Var[W_a] = \sigma^2 \sum_{i=1}^n a_i^2 \geq \frac{(\sum_{i=1}^n a_i)^2}{n} \sigma^2 = \frac{1}{n} \sigma^2$   
 $Var[\bar{Y}] = Var[\frac{1}{n} \sum_{i=1}^n Y_i] = \frac{1}{n^2} \sum_{i=1}^n Var[Y_i] = \frac{1}{n^2} n \sigma^2 = \frac{1}{n} \sigma^2$   
 $\Rightarrow Var[W_a] \geq Var[\bar{Y}]$

## 5 Problem 5

Let  $\bar{Y} = \frac{1}{n} \sum_i X_i$  where  $E[X_i] = \mu$  and  $Var[X_i] = \sigma^2$

Consider:

$$W_1 = [\frac{n-1}{n}] \bar{Y}$$

$$W_2 = \frac{\bar{Y}}{2}$$

i. Show  $W_1$  and  $W_2$  are both biased estimators, and take limit  $n \rightarrow \infty$   
 $E[W_1] = E[[\frac{n-1}{n}] \bar{Y}] = [\frac{n-1}{n}] E[\bar{Y}]$   
 $= [\frac{n-1}{n}] \mu \neq \mu$   
 $\Rightarrow \mu - E[W_1] = \frac{1}{n} \mu$   
 $\Rightarrow \lim_{n \rightarrow \infty} (\mu - E[W_1]) = 0$   
 $W_1$  approaches an unbiased estimator of  $\mu$  as the sample size becomes large  
 $E[W_2] = E[\frac{\bar{Y}}{2}] = \frac{E[\bar{Y}]}{2}$   
 $= \frac{\mu}{2} \neq \mu$   
 $\Rightarrow \mu - E[W_2] = \frac{1}{2} \mu$   
 $\Rightarrow \lim_{n \rightarrow \infty} (\mu - E[W_2]) = \frac{\mu}{2}$

As sample size gets large,  $W_2$  remains a biased estimator of  $\mu$  with constant bias

ii. Find probability limits of  $W_1$  and  $W_2$   
 $\text{plim}_{n \rightarrow \infty} W_1 = \lim_{n \rightarrow \infty} E[W_1] = \lim_{n \rightarrow \infty} \frac{n-1}{n} \mu = \mu$   
 $\text{plim}_{n \rightarrow \infty} W_2 = \lim_{n \rightarrow \infty} E[W_2] = \frac{\mu}{2}$   
 $\Rightarrow W_1$  is a consistent estimator of  $\mu$ , while  $W_2$  is not

iii. Find  $Var[W_1]$  and  $Var[W_2]$

$$Var[W_1] = Var\left[\left[\frac{n-1}{n}\right]\bar{Y}\right] = \left[\frac{n-1}{n}\right]^2 Var[\bar{Y}]$$

$$= \left[\frac{n-1}{n}\right]^2 \frac{\sigma^2}{n}$$

$$Var[W_2] = Var\left[\frac{\bar{Y}}{2}\right]$$

$$= \frac{1}{4} \frac{\sigma^2}{n}$$

iv. Consider when  $\mu$  is close to 0

$$\lim_{\mu \rightarrow 0} E[W_1] = 0 = \lim_{\mu \rightarrow 0} E[\bar{Y}]$$

$$Var[W_1] = \left[\frac{n-1}{n}\right]^2 \frac{\sigma^2}{n} = \left[\frac{n-1}{n}\right]^2 Var[\bar{Y}] < Var[\bar{Y}]$$

$\Rightarrow$  For finite number of samples, as  $\mu$  approaches 0, the expected values of  $W_1$  and  $\bar{Y}$  both approach 0, but the variance of  $W_1$  is bounded above by variance of  $\bar{Y}$ . Therefore  $W_1$  is a better estimator of  $\mu$  than  $\bar{Y}$

## 6 Problem 6

$X, Y > 0$  and  $E[Y|X] = \theta X$

i. Define  $Z = \frac{Y}{X}$ , show  $E[Z] = \theta$

Let  $a(X) = \frac{1}{X}$  and  $b(X) = 0$ , by property (1)  $\Rightarrow E[Z|X] = E\left[\frac{Y}{X}|X\right] = \frac{1}{X} E[Y|X]$

$$= \frac{1}{X} \theta X = \theta$$

By property (2)  $\Rightarrow E[Z] = E[E[Z|X]] = E[\theta] = \theta$

ii. Define  $W_1 = \frac{1}{n} \sum_{i=1}^n \frac{Y_i}{X_i}$ , show  $W_1$  is unbiased for  $\theta$

$$E[W_1] = E\left[\frac{1}{n} \sum_{i=1}^n \frac{Y_i}{X_i}\right] = \frac{1}{n} \sum_{i=1}^n E\left[\frac{Y_i}{X_i}\right]$$

$$= \frac{1}{n} \sum_{i=1}^n \theta \quad \because (i)$$

$$= \theta \Rightarrow W_1 \text{ is an unbiased estimator of } \theta$$

iii. Define  $W_2 = \frac{\bar{Y}}{\bar{X}}$ , show  $W_2 \neq W_1$  but  $W_2$  is an unbiased estimator of  $\theta$

$$W_2 = \frac{\frac{1}{n} \sum_i Y_i}{\frac{1}{n} \sum_i X_i} = \frac{\sum_i Y_i}{\sum_i X_i} \neq W_1$$

$$E[W_2] = E\left[\frac{\sum_i Y_i}{\sum_i X_i}\right] = E\left[\frac{\sum_i \theta X_i}{\sum_i X_i}\right] = \theta E\left[\frac{\sum_i X_i}{\sum_i X_i}\right]$$

$$\theta$$

$W_2$  is also unbiased estimator of  $\theta$

## 7 Problem 7

Consider  $Y$ , a Bernoulli random variable  $0 < \theta < 1$ , let  $\gamma = \frac{\theta}{1-\theta}$  For  $\{Y_i | 1 \leq i \leq n\}$ , define  $G = \frac{\bar{Y}}{1-\bar{Y}}$

i. Why is  $G$  not an unbiased estimator of  $\gamma$

$$\begin{aligned} E[\gamma - G] &= E\left[\frac{\theta}{1-\theta} - \frac{\bar{Y}}{1-\bar{Y}}\right] = E\left[\frac{\theta(1-\bar{Y}) - \bar{Y}(1-\theta)}{(1-\theta)(1-\bar{Y})}\right] \\ &= E\left[\frac{\theta - \bar{Y}}{(1-\theta)(1-\bar{Y})}\right] \end{aligned} \tag{a}$$

$$= E\left[\frac{\theta}{(1-\theta)(1-\bar{Y})}\right] - E\left[\frac{\bar{Y}}{(1-\theta)(1-\bar{Y})}\right] \tag{b}$$

$$\Rightarrow \lim_{\theta \rightarrow 0+} E[\gamma - G] = E\left[\frac{\bar{Y}}{1-\bar{Y}}\right] \geq 0$$

$\Rightarrow G$  is not an unbiased estimator of  $\gamma$

ii. Show that  $G$  is a consistent estimator of  $\gamma$

$$\text{Let } X_n = \bar{Y}_n = \frac{1}{n} \sum_i Y_i \Rightarrow \text{plim}(X_n) = \theta$$

$$\text{Let } Z_n = 1 - \bar{Y}_n = 1 - \frac{1}{n} \sum_i Y_i \Rightarrow \text{plim}(Z_n) = 1 - \theta$$

$$\because \text{plim}(X_n) = \alpha \text{ and } \text{plim}(Z_n) = \beta \Rightarrow \text{plim}\left(\frac{X_n}{Z_n}\right) = \frac{\alpha}{\beta}$$

$$\Rightarrow \text{plim}(G) = \frac{\theta}{1-\theta}$$

$\Rightarrow G$  is a consistent estimator of  $\gamma$

## 8 Problem 8

Consider the survey as a Bernoulli trial with  $p = 0.65$ ,  $q = 0.35$ ,  $n = 200$

i. Find  $E[X]$

$$E[X] = np = 130$$

ii. Find  $\sigma[X]$

$$\sigma[X] = \sqrt{npq} = 6.75$$

iii. Only 115 people of the sample voted yes out of 200. Since sample count is relatively large, we can approximate the results with a z-test using central limit theorem

$$Z = \frac{130-115}{6.75} = 2.22$$

The number of people who voted in the sampled population is about 2.22 standard deviations below expected  $\Rightarrow P[X \leq 115] = P[X \leq \mu - 2.2\sigma] = 1.39\%$

iv. Explain results to a layperson

If the initial claims of 65% support for the dictator were true, then we would expect a

sample of the population to reflect the claims. However, based on the survey, the outcome deviates greatly from the expectation. In fact, the likelihood that there are 115 or fewer people supporting the dictator in a sample of 200 is only 1.39%. Therefore with a very high confidence (98.61%) we can reject the dictator's claim that the support for his rule is 65%