Tutorial 3 Solutions

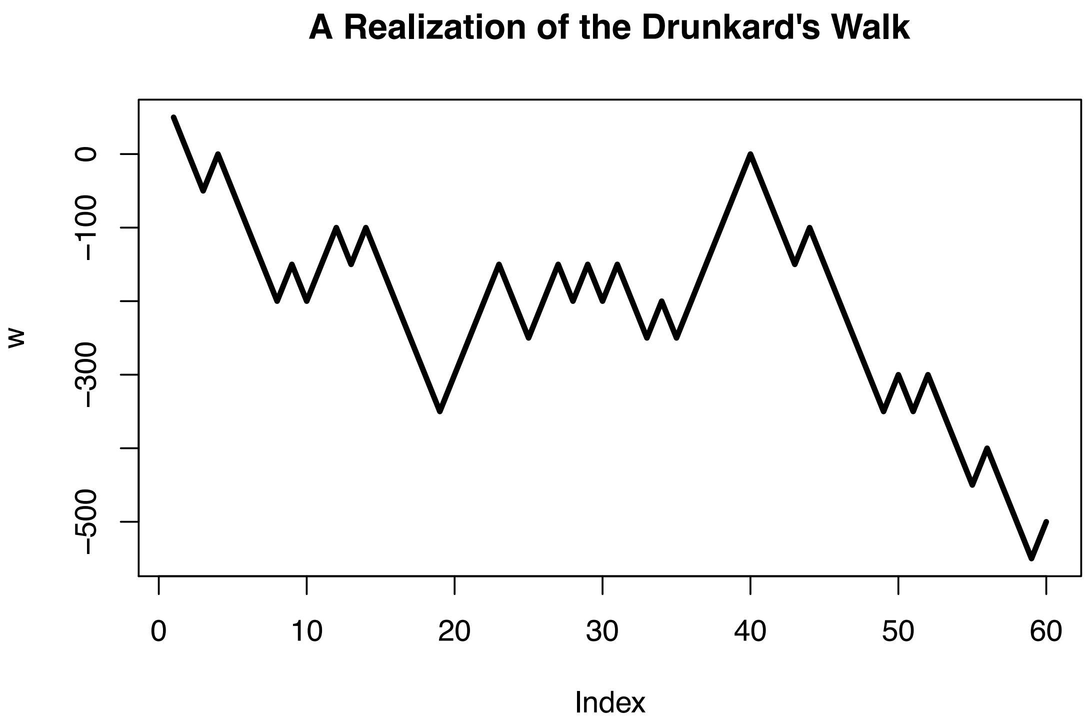
STAT 3013/4027/8027

---

1. SI: 2.6, 2.8, 2.10 **Ans.** See the handwritten pages.

- **Question 2: R code:**

```
##
set.seed(10)
x <- rbinom(60, 1, 0.5)
x[x==0] <- -1
x <- x*50

w <- cumsum(x)

##
plot(w, type="l", lwd=3, main="A Realization of the Drunkard's Walk")
```




```
## Let's examine the sampling distirbution of W
set.seed(10)
S <- 10000
W <- rep(0, S)

for(s in 1:S){
```

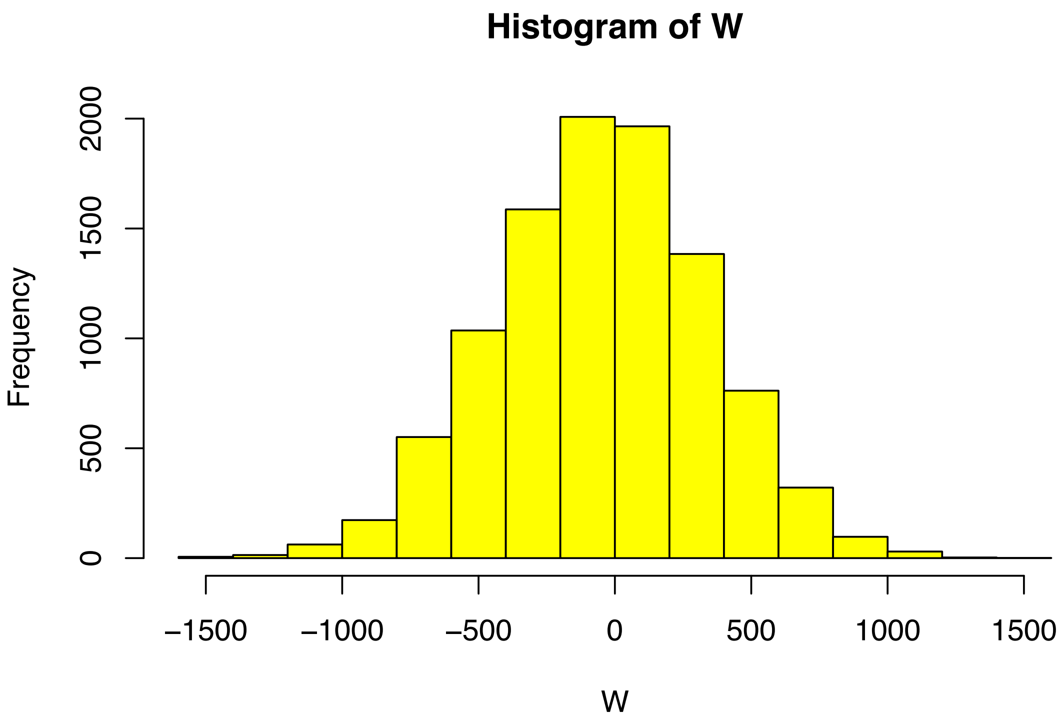
```r
  x <- rbinom(60, 1, 0.5)
  x[x==0] <- -1
  x <- x*50
  W[s] <- sum(x)
}

##
mean(W)
```

```
## [1] 1.13
```

```r
var(W)
```

```
## [1] 150870.8
```

```r
##
hist(W, col="yellow")
```



- **Question 3**

Let's first work out the exact answer.

$$
\begin{aligned}
\int_0^1 cos(2\pi x)dx &= \frac{1}{2\pi} sin(2\pi x)\Big|_0^1 \\
&= 0
\end{aligned}
$$

- Note. Suppose we consider $U \sim$ Uniform(0,1). Let's look at the $E[U]$.

$$E[U] = \int_0^1 u f(u) du = \int_0^1 u \ du$$

- Now let's look at the expected value of the function: $cos(2\pi u)$:

$$E[g(U)] = \int_0^1 cos(2\pi u) \ du$$

So our integral of interest is the expected value of the function. We can approximate that via simulation.

$$\hat{I}(g(u)) = \frac{1}{S} \sum_1^S cos(2\pi u)$$

```
set.seed(1001)
S1 <- 100

u <- runif(S1)
g.u <- cos(2*pi*u)
mean(g.u)
```

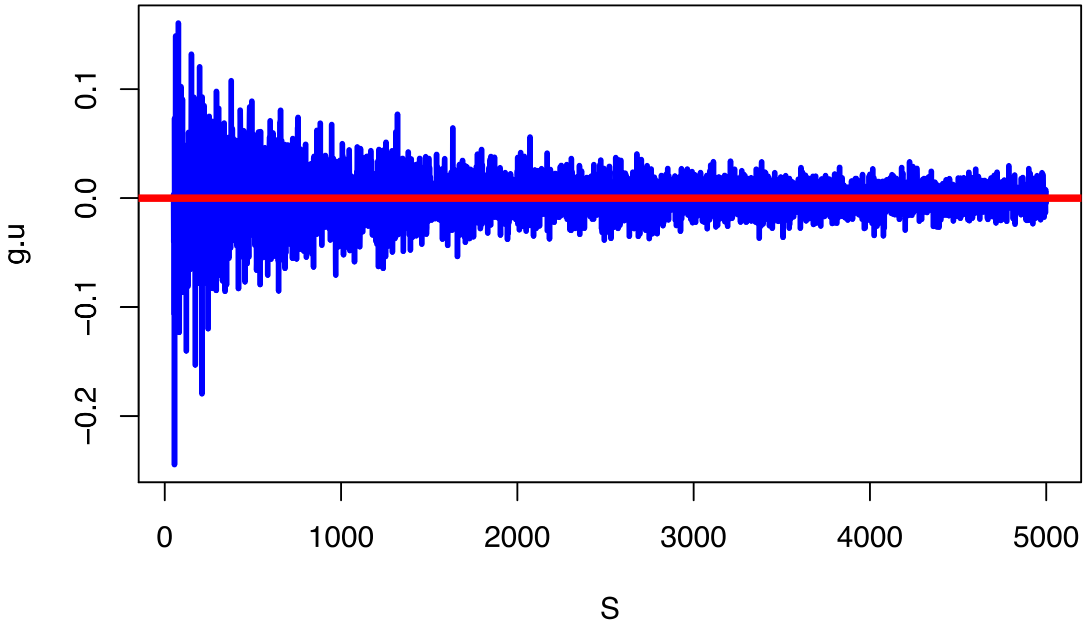
```
## [1] 0.05382692
```

```
set.seed(1001)
S2 <- 1000

u <- runif(S2)
g.u <- cos(2*pi*u)
mean(g.u)
```

```
## [1] 0.006431414
```

- Now let's plot for increasing values of S:

```
S <- 50:5000
g.u <- rep(0, length(S))

c <- 1
for(s in S){
g.u[c] <- mean(cos(2*pi* runif(s)))
c <- c+1
}

plot(S, g.u, type="l", lwd=3, col="blue")
abline(h=0, col="red", lwd=4)
```



- As we can work out the $E[g(u)]$ and the $V[g(u)]$ we could also use the CLT theorem for calculations based on the $\frac{1}{S}\sum_1^S cos(2\pi u)$.
- **Question 4**
- For this question we can't work out an analytical solution (some type of approximation must be performed).

$$
\begin{aligned}
E[I(g(u))] &= \int_0^1 cos(2\pi u^2)du \\
&\approx \frac{1}{S}\sum_1^S cos(2\pi u^2)
\end{aligned}
$$

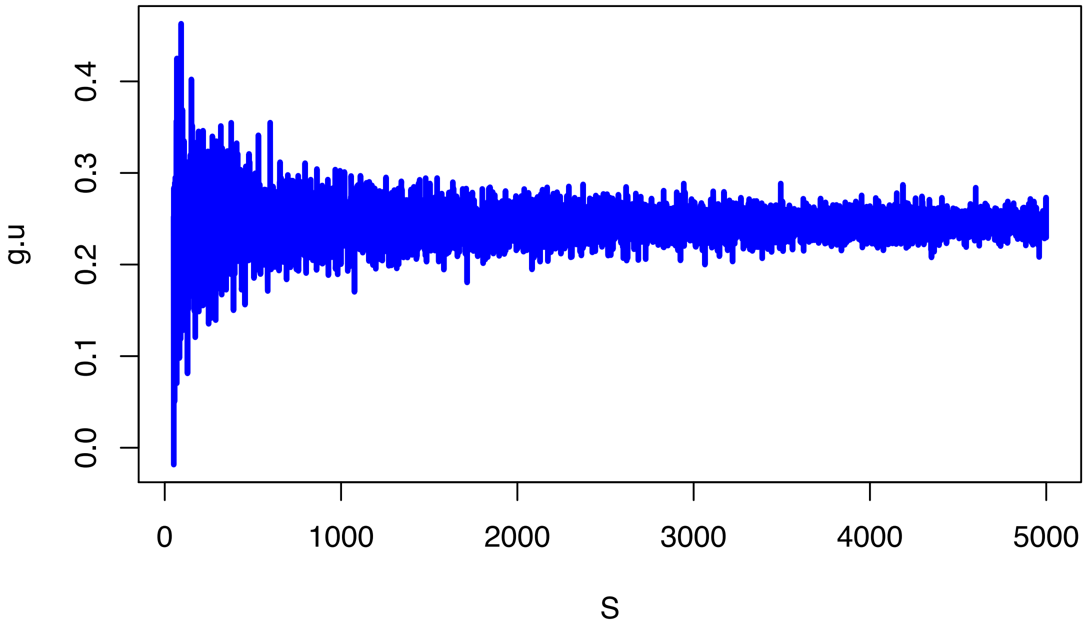
```
set.seed(1001)
S1 <- 100

u <- runif(S1)
g.u <- cos(2*pi*u^2)
mean(g.u)
```

```
## [1] 0.3031969
```

```
set.seed(1001)
S2 <- 1000

u <- runif(S2)
g.u <- cos(2*pi*u^2)
mean(g.u)
```

```
## [1] 0.2592541
```

```r
S <- 50:5000
g.u <- rep(0, length(S))

c <- 1
for(s in S){
g.u[c] <- mean(cos(2*pi* runif(s)^2))
c <- c+1
}

plot(S, g.u, type="l", lwd=3, col="blue")
```



GJJ 2.6 $\quad X_1, \ldots, X_n \sim \text{Bernoulli}(\theta)$

$$f(x;\theta) = \theta^{x}(1-\theta)^{1-x}$$

CR lower bound for estimators of θ:

$$V(\hat{\theta}) \geq I_\theta^{-1} \; ; \quad I_\theta = -E\left(\frac{\partial^2 \ell}{\partial \theta^2}\right)$$

$$= E\left(\left(\frac{\partial \ell}{\partial \theta}\right)^2\right)$$

$$\Rightarrow L(\theta) = \prod_{i=1}^{n} \theta^{x_i}(1-\theta)^{1-x_i}$$

$$= \theta^{\sum x_i}(1-\theta)^{n-\sum x_i}$$

$$\ell(\theta) = \sum x_i \log(\theta) + (n-\sum x_i)\log(1-\theta)$$

$$\frac{\partial \ell}{\partial \theta} = \frac{\sum x_i}{\theta} - (n-\sum x_i)\frac{1}{1-\theta}$$

$$\frac{\partial^2 \ell}{\partial \theta} = \frac{-\sum x_i}{\theta^2} - \frac{(n-\sum x_i)}{(1-\theta)^2}$$

$$\Rightarrow E\left(\frac{\partial^2 \ell}{\partial \theta}\right) = -\frac{E(\sum x_i)}{\theta^2} - \frac{E((n-\sum x_i))}{(1-\theta)^2}$$

$$E\left(\sum x_i\right) = \sum E(x_i) = \sum \theta = n\theta$$

$$= -\frac{n\theta}{\theta^2} - \frac{n-n\theta}{(1-\theta)^2} = \frac{-n\cancel{\theta}}{\theta^{\cancel{2}}} - \frac{n\cancel{(1-\theta)}}{(1-\theta)^{\cancel{2}}}$$

$$= \frac{-n}{\theta} - \frac{n}{(1-\theta)} = \frac{-n(1-\theta) - n\theta}{\theta(1-\theta)}$$

$$= \frac{-n}{\theta(1-\theta)} \quad \therefore \quad I_\theta = -E\left(\frac{\partial^2 \ell}{\partial \theta^2}\right)$$

$$= -\left(\frac{n}{\theta(1-\theta)}\right)$$

$$= \frac{n}{\theta(1-\theta)}$$

$$C\text{-}R\ LB = \frac{\theta(1-\theta)}{n}$$

- Now the C-R LB for an estimator of $\theta^2 = g(\theta)$:

$$V(\hat{\theta}^2) \leq \left(\left(\frac{\partial g}{\partial \theta}\right)^2 I_\theta^{-1}\right)$$

$$= \frac{(2\theta)^2 \ \theta(1-\theta)}{n}$$

$$= \frac{4\theta^3(1-\theta)}{n}$$

GJJ Q 2.8.) $X_1, \dots, X_n \overset{iid}{\sim} N(\mu, \theta = \sigma^2)$; assume μ is known.

For the CRLB does this make a difference in this case? Why or why not?

$$S^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1} \text{ ; Note: } W = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2_{n-1}$$

$$E(W) = n-1$$
$$V(W) = 2(n-1)$$

$$E\left(\frac{(n-1)}{\sigma^2} S^2\right) = \frac{(n-1)}{\sigma^2} E(S^2) = (n-1)$$

$$E(S^2) = \sigma^2 \therefore S^2 \text{ is an unbiased estimator of } \sigma^2$$

$$V\left(\frac{(n-1)}{\sigma^2} S^2\right) = \frac{(n-1)^2}{(\sigma^2)^2} V(S^2) = 2(n-1)$$

$$V(S^2) = \frac{2(\sigma^2)^2}{(n-1)}$$

- Now let's determine the CRLB:

$$L(\theta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2\sigma^2}(x_i-\mu)^2\right) = \prod \frac{1}{\sqrt{2\pi\theta}} \exp\left(-\frac{1}{2\theta}(x_i-\mu)^2\right)$$

$$= (2\pi)^{-n/2} \theta^{-n/2} \exp\left(-\frac{1}{2\theta} \sum (x_i-\mu)^2\right)$$

$$\ell(\theta) = -n/2 \log(2\pi) - n/2 \log(\theta) - \frac{1}{2\theta} \sum (x_i - \mu)^2$$

$$\Rightarrow \ell'(\theta) = -\frac{n}{2}\frac{1}{\theta} + \frac{1}{2\theta^2} \sum (x_i-\mu)^2$$

$$\ell''(\theta) = \frac{n}{2}\frac{1}{\theta^2} - \frac{1}{\theta^3} \sum (x_i-\mu)^2$$

$$-E(\ell''(\theta)) = -\frac{n}{2}\frac{1}{\theta^2} + \frac{1}{\theta^3} E\left(\sum (x_i-\mu)^2\right)$$

$$= -\frac{n}{2}\frac{1}{\theta^2} + \frac{1}{\theta^3} \underbrace{\sum \underbrace{E((x_i-\mu)^2)}_{=\theta}}_{=n\theta}$$

$$= -\frac{n}{2}\frac{1}{\theta^2} + \frac{2n}{2\theta^2} = \frac{n}{2\theta^2} \Rightarrow CRLB = \frac{2\theta^2}{n}$$

- For finite n: $CRLB = \frac{2\theta^2}{n} = \frac{2\sigma^4}{n} < \frac{2\sigma^4}{n-1} = V(S^2)$

However, as $n \to \infty$ they both go to $\underline{\underline{0}}$.

- Now we want to know the value 'c' which minimizes the MSE of:

$$c \sum_{i=1}^{n} (x_i - \bar{x})^2$$

- We know $MSE(\hat{\theta}) = V(\hat{\theta}) + [Bias(\hat{\theta})]^2$

$$\Rightarrow V\left(c \sum (x_i - \bar{x})^2\right) = V\left(c \frac{(n-1)}{(n-1)} \sum (x_i - \bar{x})^2\right)$$

$$= V\left(c(n-1)S^2\right) = c^2 (n-1)^2 \frac{2\sigma^4}{(n-1)}$$

$$= c^2 (n-1) 2\sigma^4$$

$$\Rightarrow E\left(c \sum (x_i - \bar{x})^2\right) = E\left(c(n-1)S^2\right)$$

$$= c(n-1)E(S^2) = c(n-1)\sigma^2$$

$$\therefore MSE\left(c \sum (x_i - \bar{x})^2\right) = c^2(n-1)2\sigma^4 + \left[c(n-1)\sigma^2 - \sigma^2\right]^2$$

minimize MSE wrt c

$$= c^2(n-1)2\sigma^4 + c^2(n-1)^2\sigma^4 - 2c(n-1)\sigma^4 + \sigma^4$$

$$\frac{d}{dc} = 2c(n-1)2\sigma^4 + 2c(n-1)^2\sigma^4 - 2(n-1)\sigma^4 = 0$$

$$= (n-1)\sigma^4 \left[4c + 2c(n-1) - 2\right] = 0$$

$$\Rightarrow 4c + 2cn - 2c - 2 = 0$$

$$\Rightarrow c + cn - 1 = 0$$

$$\Rightarrow c(1+n) = 1 \Rightarrow c = \frac{1}{n+1}$$

Check to see that we have a minimum

$$\frac{d^2}{dc^2} = (n-1)\sigma^4\left[4 + 2(n-1)\right] > 0$$

for $\underline{n \geq 2}$

$\therefore$ we have a minimum.

GJJ Q 2.10) $X_1, \ldots, X_n \overset{iid}{\sim} N(\theta, \sigma^2)$, where σ^2 is known.

Lemma 2.1] Under our standard regularity conditions we have:

$$\underline{I} = -E\left(\frac{\partial^2 \ell}{\partial \theta^2}\right)$$

$$L(\theta) = \prod_{i=1}^{n} (2\pi\sigma^2)^{-1/2} \exp\left(-\frac{1}{2\sigma^2}(x_i - \theta)^2\right)$$

$$= (2\pi\sigma^2)^{-n/2} \exp\left(-\frac{1}{2\sigma^2}\Sigma(x_i - \theta)^2\right)$$

$$\ell(\theta) = -n/2 \log(2\pi\sigma^2) - 1/2\sigma^2 \,\Sigma(x_i - \theta)^2$$

$$\ell'(\theta) = +\frac{2}{2\sigma^2}\Sigma(x_i - \theta) = \frac{1}{\sigma^2}\left[\Sigma x_i - n\theta\right]$$

$$\ell''(\theta) = -\frac{n}{\sigma^2} \Rightarrow I = -E(-n/\sigma^2) = \frac{n}{\sigma^2}$$

$$\therefore \ CRLB = 1/I = \sigma^2/n$$