Problem 1

a) Let $Y = \max\{X_1, \cdots, X_n\}$ $\qquad X_i \overset{iid}{\sim} N(\theta, 1)$

then $F_Y(x_{(n)} | \theta) = \prod_{i=1}^{n} P(X_i \leq x_{(n)}) = \Phi_\theta(x_{(n)})^n$

↳ CDF of $N(\theta,1)$

$f_Y(x_{(n)} | \theta) = \frac{d}{dx_{(n)}} \Phi_\theta(x_{(n)})^n = n \cdot \phi_\theta(x_{(n)} | \theta) \cdot \Phi_\theta(x_{(n)})^{n-1}$

↳ pdf of $N(\theta,1)$

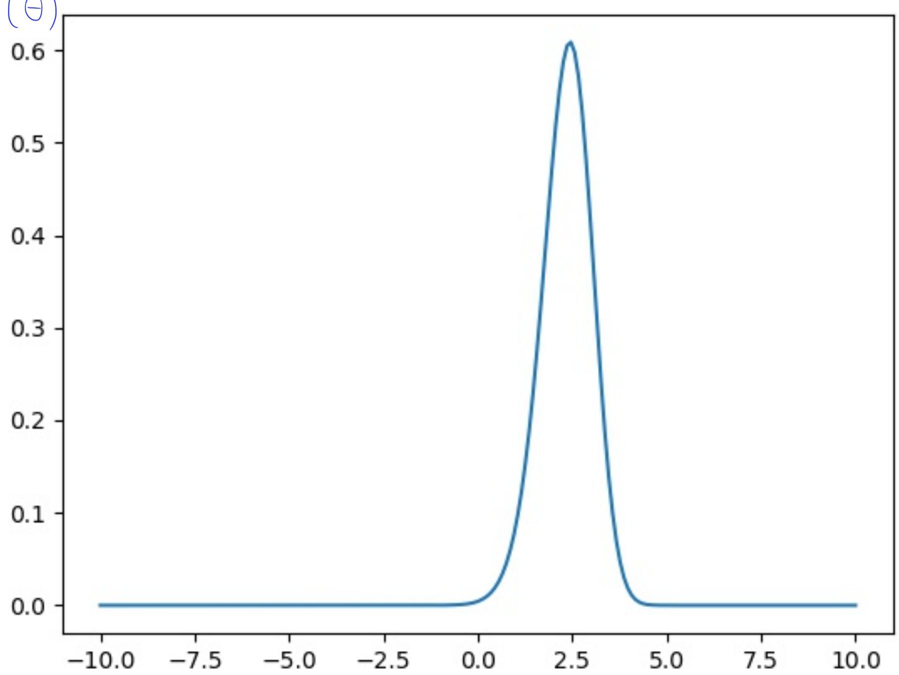
$\therefore\ L(\theta | x_{(n)}) = f_Y(x_{(n)} | \theta) = n \cdot \exp\left\{-\frac{(x_{(n)} - \theta)^2}{2}\right\} \cdot \left[\int_{-\infty}^{x_{(n)}} \exp\left\{-\frac{(x-\theta)^2}{2}\right\} d\theta\right]^{n-1}$

by plot, $\hat{\theta}_{MLE} = 2.44$

At $X_{(n)} = 3.5$. $\quad n = 5$.

$L(\theta | X_{(n)})$ vs $\hat{\theta}$.

$L(\theta)$



$\hat{\theta}$

b). $\bar{Y} = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad x_i \overset{iid}{\sim} N(\theta, 1)$

then $\bar{Y} \sim N(\theta, \frac{1}{n})$

$\therefore L(\theta \mid \bar{Y}=\bar{y}) = f_{\bar{Y}}(\bar{y}\mid\theta) = \frac{1}{\sqrt{2\pi\frac{1}{n}}} \exp\left\{-\frac{(\bar{y}-\theta)^2}{2/n}\right\}$

by plot. $\hat{\theta}_{MLE} = 3.1$

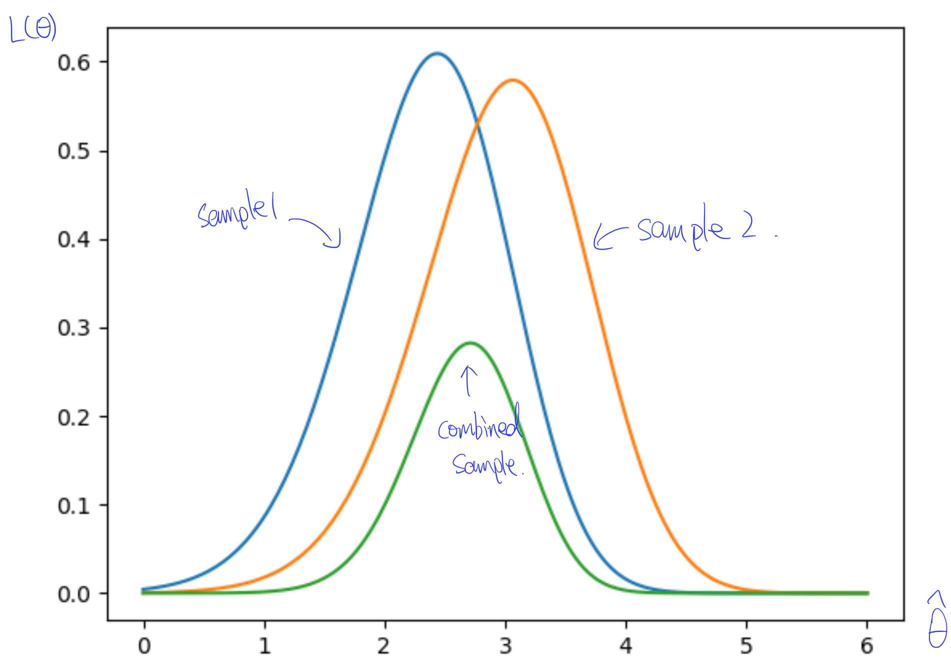
Likelyhood for both sample.

$L(\theta \mid Y=x_{(n)}, \bar{Y}=\bar{y}) = f_{\bar{Y}}(\bar{y}\mid\theta)\cdot f_Y(x_{(n)}\mid\theta)$ As independent.

$= n\cdot \phi_\theta(x_{(n)}\mid\theta)\cdot\Phi_\theta(x_{(n)})^{n-1}\cdot \frac{1}{\sqrt{2\pi\frac{1}{n}}}\exp\left\{-\frac{(\bar{y}-\theta)^2}{2/n}\right\}$

$\hat{\theta}_{MLE,\,combine} = 2.7$

$L(\theta)$ vs $\hat{\theta}$



Problem 2.

a). marginal for θ_1

$$g(\theta_1 | x_1, \cdots, x_m) = \frac{f(x_1, \cdots, x_m | \theta_1) \cdot \pi(\theta_1)}{\int_0^\infty f(x_1, \cdots, x_m | \theta_1) \cdot \pi(\theta_1)\, d\theta_1}$$

$$= \frac{\prod_{i=1}^{m} \frac{e^{-\theta_1} \theta_1^{x_i}}{x_i!} \cdot \frac{1^2}{\Gamma(2)} \theta_1 e^{\theta_1}}{\int_0^\infty \frac{e^{-m\theta_1} \theta_1^{\Sigma x_i}}{\prod_{i=1}^{m}(x_i!)} \cdot \theta_1 e^{\theta_1}\, d\theta_1} \begin{matrix} \leftarrow (1) \\ \leftarrow (2). \end{matrix}$$

$$(2) = \frac{1}{\prod_{i=1}^{m}(x_i!)} \int_0^\infty e^{-\theta_1(m+1)} \cdot \theta_1^{\Sigma x_i + 1}\, d\theta_1 = \frac{\Gamma(\Sigma x_i + 2)}{\prod_{i=1}^{m}(x_i!)\,(1+m)^{(\Sigma x_i+2)}} \underbrace{\int_0^\infty \frac{(1+m)^{(\Sigma x_i+2)}}{\Gamma(\Sigma x_i+2)} \cdot \theta_1^{\Sigma x_i+1} \cdot e^{-\theta_1(m+1)}\, d\theta_1}_{Gamma(\Sigma x_i + 2,\ m+1)}$$

$$= \frac{\Gamma(\Sigma x_i + 2)}{\prod_{i=1}^{m}(x_i!)\,(1+m)^{(\Sigma x_i+2)}}$$

$$\frac{(1)}{(2)} = \frac{e^{-\theta_1(m+1)}\ \theta_1^{\Sigma x_i+1}}{\cancel{\prod_{i=1}^{m}(x_i!)}} \cdot \frac{\cancel{\prod_{i=1}^{m}(x_i!)}\ (1+m)^{(\Sigma x_i+2)}}{\Gamma(\Sigma x_i+2)}$$

$$= \frac{(1+m)^{(\Sigma x_i+2)}}{\Gamma(\Sigma x_i+2)}\, e^{-\theta_1(m+1)}\, \theta_1^{\Sigma x_i+1}$$

pdf for $Gamma(\Sigma x_i + 2,\ m+1)$

$\therefore$ marginal posterior distn of θ_1 is $Gamma(\sum_{i=1}^{m} x_i + 2,\ m+1)$

Similarly: marginal posterior distn of θ_2 is $Gamma(\sum_{i=1}^{n} Y_i + 2,\ n+1)$

where $g(\theta_2 | Y_1, \cdots Y_n) = \frac{(1+n)^{\Sigma Y_i + 2}}{\Gamma(\Sigma Y_i + 2)}\, e^{-\theta_2(n+1)}\, \theta_2^{\Sigma Y_i + 1}$.

joint posterior for θ_1, θ_2.

$$g(\theta_1, \theta_2 \mid X_1, \cdots, X_m, Y_1, \cdots, Y_n) = g_{\theta_1}(\theta_1 \mid X_1, \cdots, X_m) \cdot g_{\theta_2}(\theta_2 \mid Y_1, \cdots, Y_n)$$

$$= \frac{(1+m)^{(\Sigma X_i + 2)}}{\Gamma(\Sigma X_i + 2)} e^{-\theta_1(m+1)} \theta_1^{\Sigma X_i + 1} \frac{(1+n)^{\Sigma Y_i + 2}}{\Gamma(\Sigma Y_i + 2)} e^{-\theta_2(n+1)} \theta_2^{\Sigma Y_i + 1}.$$

b).

$$\text{mean}(\theta_1) = \frac{\alpha}{\beta} = \frac{\sum_{i=1}^{m} X_i + 2}{m+1} = \frac{219}{112} = 1.955$$

$$\text{mean}(\theta_2) = \frac{\alpha}{\beta} = \frac{\sum_{i=1}^{n} Y_i + 2}{n+1} = \frac{68}{45} = 1.511$$

$$\text{mode}(\theta_1) = \frac{\alpha - 1}{\beta} = \frac{\sum_{i=1}^{m} X_i + 1}{m+1} = 1.946$$

$$\text{mode}(\theta_2) = \frac{\alpha - 1}{\beta} = \frac{\sum_{i=1}^{n} Y_i + 1}{n+1} = 1.489$$

R code.

95% CI (θ_1) = [qgamma(0.025, 219, 112), qgamma(0.975, 219, 112)]

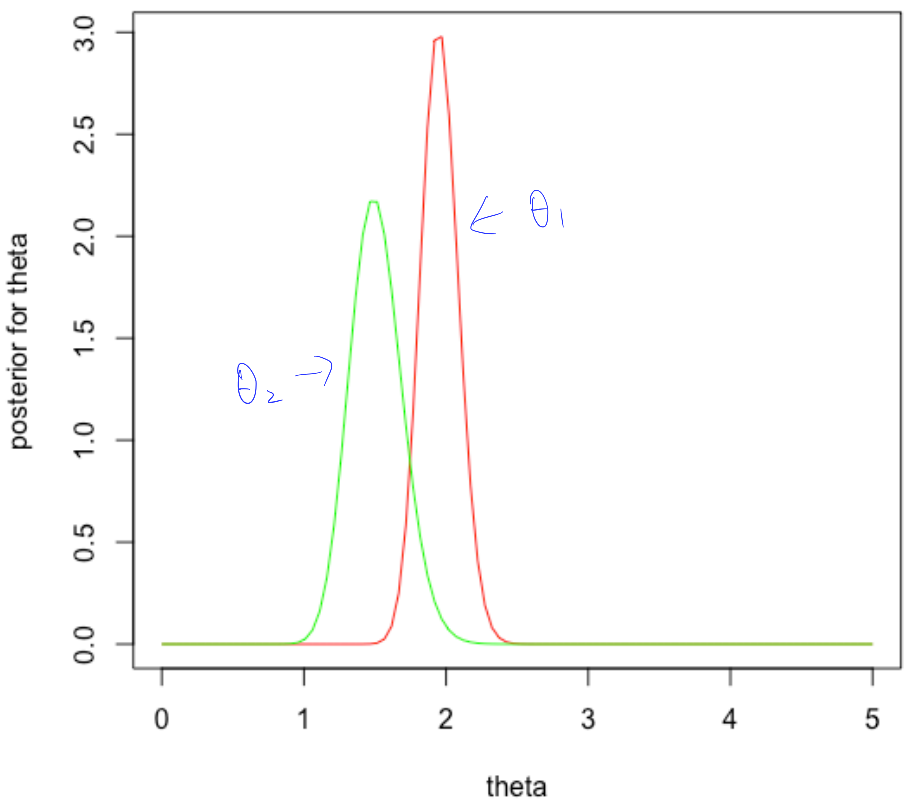
= [1.7049, 2.2227]

95% CI (θ_2) = [qgamma(0.025, 68, 45), qgamma(0.975, 68, 45)]

[1.1734, 1.8908]

c)

posterier vs θ.



θ_1 has higher confidence on its maximum. since it has higher number of sample

d) $$\Pr(\Theta_1 > \Theta_2) = \int_0^\infty \Pr(\Theta_2 < \theta_1 \mid \Theta_1 = \theta_1) \cdot \Pr(\Theta_1 = \theta_1)\, d\theta_1$$

$$= \int_0^\infty F_{\theta_2}(\theta_1) \cdot f_{\theta_1}(\theta_1)\, d\theta_1$$

$$= E_{\theta_1}[F_{\theta_2}(\theta_1)]$$

$$= 0.973$$ (in 10000 sample, use Monte Carlo in R)

sample from $\theta_1 \sim Gamma(219, 112)$

$p = mean(pgamma(\theta_1, 68, 45))$.

there is 97.3% Prob $\theta_1 > \theta_2$.

thus. we conclude that we are 97.3% confidence less than bachelor have more children than Bachelor's or higher.

e). $$P(X^f \mid x_1, \cdots, x_m) = \int_0^\infty \underbrace{P(X^f \mid X_{obs}, \theta_1)}_{\sim Poi(\theta_1)} \cdot \underbrace{\pi(\theta_1 \mid X_{obs})}_{Gamma(219,112)} d\theta_1$$

$$= \int_0^\infty \frac{\theta_1^{x^f} e^{-\theta_1}}{x^f!} \cdot \frac{112^{219}}{\Gamma(219)} \cdot \theta_1^{219-1} \cdot e^{-(112)\theta_1}\, d\theta_1$$

$$= \frac{112^{219}}{\Gamma(x^f+1)\Gamma(219)} \int_0^\infty \theta_1^{x^f+219-1} \cdot e^{-(112+1)\theta_1}\, d\theta_1$$

$$= \frac{112^{219}}{\Gamma(x^f+1)\Gamma(219)} \int_0^\infty \underbrace{\theta_1^{x^f+219-1} \cdot e^{-\theta_1}}_{dv} \underbrace{e^{-112\theta_1}}_{u} d\theta_1$$

pmf for. $NB\left(\underset{\alpha}{219}, \underset{\frac{1}{\beta+1}}{\frac{112}{112+1}}\right)$

$$= \frac{112^{219}}{\Gamma(x^f+1)\Gamma(219)} \frac{\Gamma(x^f+219)}{(112+1)^{x^f+219}}$$

$$= \binom{x^f+219-1}{x^f} \cdot \left(\frac{112}{112+1}\right)^{219} \cdot \left(\frac{1}{112+1}\right)^{x^f}$$

Thus: we have: $X^f | x_{obs} \sim NB\left(219, \frac{112}{112+1}\right)$

Similarly: we have $Y^f | y_{obs} \sim NB\left(68, \frac{45}{45+1}\right)$

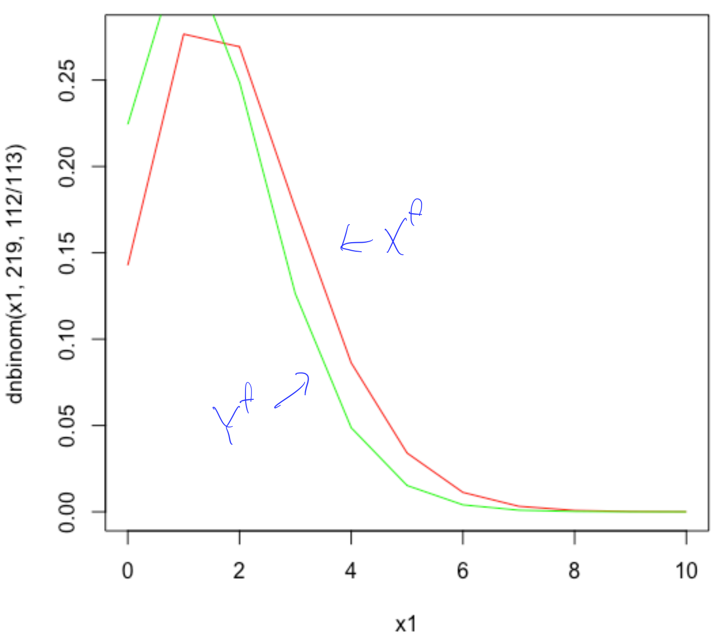
we calculate $P_r(Y^f < X^f) = \sum_{x^f=0}^{\infty} P_r(Y^f < x^f) \cdot P_r(X^f = x^f)$

$= E_{X^f}\left[F_{Y^f}(x^f - 1) \right]$

$= 0.47$

we are 47% confidence that $X^f > Y^f$

f).



Compare event $\{X^f > Y^f\}$ their dist^n are closer together. ie: there are more Area in common. But for event $\{\theta_1 > \theta_2\}$. they are more separated.

Problem 3.

a) $\int_{-\infty}^{\infty} f(x)\,dx \approx \int_{-10}^{10} 20 f(x) \cdot \frac{1}{20}\,dx$

$$= E_{U[-10,10]}(20f(u)) + \alpha\left(E_{U[-10,10]}(20\, g(u)) - \int_{-10}^{10} (1 - \frac{|x|}{5}\, dx.)\right)$$

$$\alpha^* = \frac{Cov\left(E_{U[-10,10]}(20f(u)),\ E_{U[-10,10]}(20\, g(u))\right)}{Var\left(E_{U[-10,10]}(20 \cdot g(u))\right)} = 0.2924 \quad \text{by sample.}$$

then $\int_{-\infty}^{\infty} f(x)\,dx \approx 2.522$ by R code.

```r
n = 50

rv = vector("numeric", n)

for (i in 1:50) {
    u = runif(500, -10, 10)
    theta = 20*exp(-u^2/2)
    theta_star = 20*(1-abs(u)/5)
    alpha = cov(theta, theta_star)/var(theta_star)
    rv[i] = mean(theta + alpha * (theta_star - 0))
}

mean(rv)

```

var(rv) = 0.1578

mean(rv) = 2.5897

b). $\int_{-\infty}^{\infty} f(x)\,dx \approx \int_{-10}^{10} 20 f(x)\cdot\frac{1}{20}\,dx.$

by CV: $\hat{\theta}_{cv} = E_{U(-10,10)}[20 f(x)] + \alpha\left(E_{U(-10,10)}(20 h(x)) - \int_{-10}^{10} h(x)\,dx\right)$

$\int_{-10}^{10} 1-\frac{x^2}{25}\,dx = -\frac{20}{3}$

$\alpha^* = \frac{Cov(\hat{\theta}_{MC}, \hat{\theta}^*)}{Var(\hat{\theta}^*)}$

$= 0.1064$ by sample.

$\hat{\theta}_{cv} = 2.5262$. $Var(\hat{\theta}_{cv}) = 0.1233$

```r
n=50

rv2 = vector("numeric", n)

for (i in 1:50) {
    u = runif(500, -10, 10)
    theta = 20*exp(-u^2/2)
    theta_star = 20*(1-u^2/25)
    alpha = cov(theta, theta_star)/var(theta_star)
    rv2[i] = mean(theta + alpha * (theta_star + 20/3))
}

mean(rv2)
var(rv2)
```

Variance are both small and similar for g(x) and h(x)

Thus there is no simple way to tell which is more efficient.

Problem 4.

a). Let $f(x) = \frac{x^2}{\sqrt{2\pi}} \exp\{-x^2/2\}$

$$\theta = \int_1^\infty \frac{f(x)}{g(x)} g(x)\, dx.$$

$$= E_g[x^2 \cdot (1-\Phi(1))]$$

by R. code:

```
x=rtruncnorm(1000,1)
y1= x^2 • (1- pnorm(1))
print (mean(y))
```

using AR Sampling talked in class

Answer is 0.4026

b). Let $u \overset{d}{\sim} h(x)$.

$$\theta = \int_1^\infty \frac{f(x)}{h(x)} \cdot h(x)\, dx$$

$$= E_h\left[\frac{x}{\sqrt{2\pi}} \cdot \exp\{-\tfrac{1}{2}\}\right]$$

by R code.

```
x= sqrt(rexp(1000, 0.5) +1)
y2= (x / √2π)· exp(-1/2)
print(mean(y))
```

Answer is 0.3708

c). by R code. $\sqrt{Var(y_1)/1000} = 0.008103$ (from $g(x)$)

$\sqrt{Var(y_2)/1000} = 0.00333$ (from $h(x)$)

because $\frac{f(x)}{h(x)}$ is less than 1, or $h(x)$ are "closer" to $f(x)$

so it tend to have smaller variance

Problem 5.

a) $\theta = \int_0^\infty x^2 \sin(\pi x) e^{-x/2} dx$. Let $y = \frac{1}{x+1} \Rightarrow x = \frac{1}{y} - 1 \Rightarrow dx = -\frac{1}{y^2} dy$

$$= \int_0^1 \underbrace{\left(\tfrac{1}{y}-1\right)^2 \sin\left(\pi \cdot \left(\tfrac{1}{y}-1\right)\right) \cdot \exp\left\{-\left(\tfrac{1}{y}-1\right)/2\right\} \cdot \left(-\tfrac{1}{y^2}\right)}_{f(x).} dy$$

$\tilde{\theta}_{CMC} = \frac{1}{n} \sum_{i=1}^{n} f(U_i)$ where $U_i \overset{iid}{\sim} U(0,1)$ $\hat{\theta}_{CMC} = 0.05949$ (10000 sample).

b). $\theta = \int_0^\infty \underbrace{2 \cdot x^2 \sin(\pi x)}_{f(x)} \cdot \underbrace{\tfrac{1}{2} e^{-\frac{x}{2}}}_{\text{pdf of } \exp(\frac{1}{2})} dx.$

$F_{exp}(x) = 1 - e^{-\frac{x}{2}}$ $u = 1 - e^{-\frac{x}{2}} \Rightarrow x = -2\ln(u)$

then: $U \sim \text{runif}(0,1)$

$$\left.\begin{aligned} X &= -2\ln(U) \\ Y &= -2\ln(1-U) \end{aligned}\right\} X, Y \text{ form } \exp(\tfrac{1}{2}) \text{ with negative Cov.}$$

$\tilde{\theta}_{AS} = \frac{\text{mean}(f(X)) + \text{mean}(f(Y))}{2}$ (both X, Y have 10000 sample).

$\hat{\theta}_{AS} = -0.05577$

c) by R code.

$Var(\tilde{\theta}_{CMC}) = 0.04906$. $Var(\tilde{\theta}_{AS}) = 0.03931$

$\therefore$ AS performs better. ($\sim 20\%$ better than CMC)

Problem 6.

a). $\theta = \int_0^{\pi/2} \cos(x) \sin(x) \exp\{-\cos^2(x)\} dx.$ Let $y = \frac{x}{\pi/2} \Rightarrow \frac{\pi}{2} y = x.$ $\frac{\pi}{2} dy = dx$

$= \int_0^1 \cos(\frac{\pi}{2}y) \sin(\frac{\pi}{2}y) \exp\{-\cos^2(\frac{\pi}{2}y)\} \cdot \frac{\pi}{2} dy.$

by R code. y = runif(100).

$\hat{\theta}_{CMC} = 0.2808$

b). $95\% \ CI(\hat{\theta}_{CMC}) = \hat{\theta}_{CMC} \pm 1.96 \cdot \frac{sd(\hat{\theta}_{CMC})}{\sqrt{100}} = [0.2448, 0.3169]$

c). Let $y = \cos^2(x)$ $dy = -\sin(2x) dx = -2\cos(x)\sin(x) \cdot dx.$

$\theta = -\int_0^1 -\frac{1}{2} \exp\{-y\} dy.$

$= -(\frac{1}{2} e^{-y}) \Big|_0^1$

$= 0.31606. \in [0.2448, 0.3169]$

within 95% CI.

d) for $n = 500$. $\hat{\theta}_{CMC} = 0.3054$ $95\% \ CI = [0.2905, 0.3204]$.

$n = 1000$ $\hat{\theta}_{CMC} = 0.3154$ $95\% \ CI = [0.3049, 0.3259]$

e). total length 0.02. $\Rightarrow 1.96 \times \frac{sd(\hat{\theta}_{CMC})}{\sqrt{n}} = 0.01.$

$sd(\hat{\theta}_{CMC}) = 0.1659 \Rightarrow n = \left(sd(\hat{\theta}_{CMC}) \cdot \frac{1.96}{0.01}\right)^2 = 1024$

need sample size ≥ 1024.

Problem 7.

a).

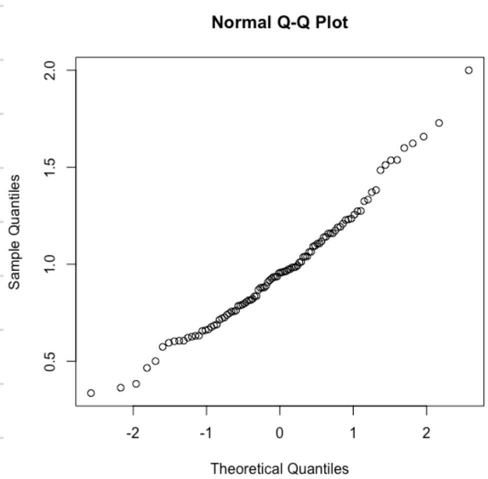
```r
temp1 = vector('numeric', 100)
temp2 = vector('numeric', 100)

for (i in 1:100) {

    n = 10

    mean1 = mean(rexp(n, 1))
    temp1[i] = mean1

    mean2 = mean(sample(c(1, 3, 9), n, TRUE, c(1/8, 1/8, 3/4)))
    temp2[i] = mean2

}
```

by R code: temp1 $\sim$ Exp(1).

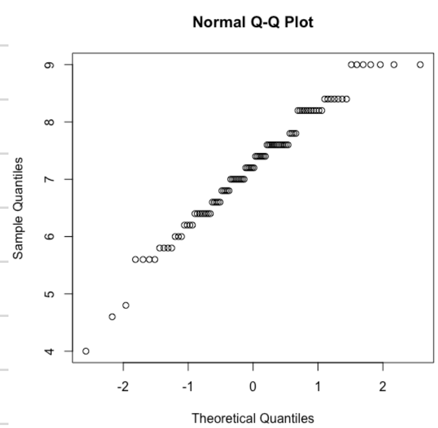
temp2 $\sim$ f(x).

b). sample mean are: mean(temp1) = 0.9712

mean(temp2) = 7.2.



qqnorm(temp1)

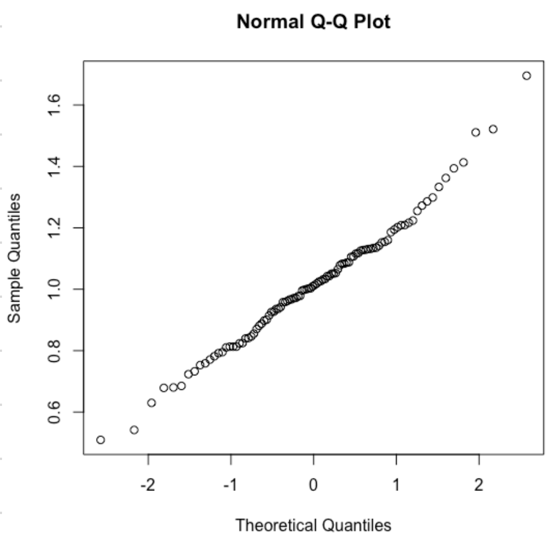


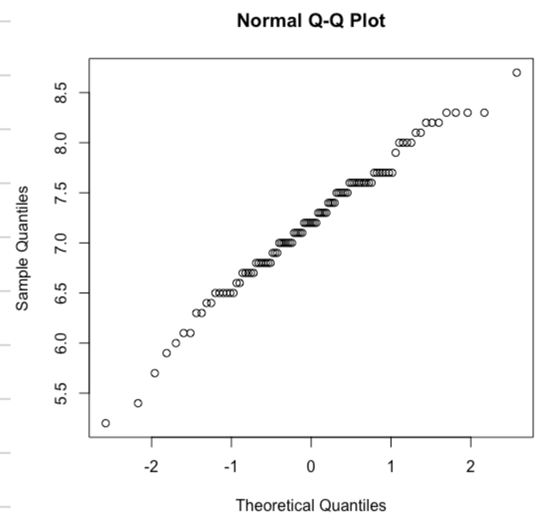
qqnorm(temp2)

As we can see in the qq-plot., the data distrubuted alone a stright line. Thus we can conclude that the mean of any distrubution is normal distrubuted, which is exactly CLT.

c). For n=20:

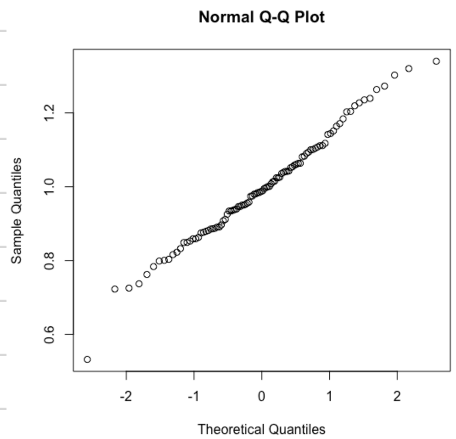
mean(temp1) = 1.014   mean(temp2) = 7.183

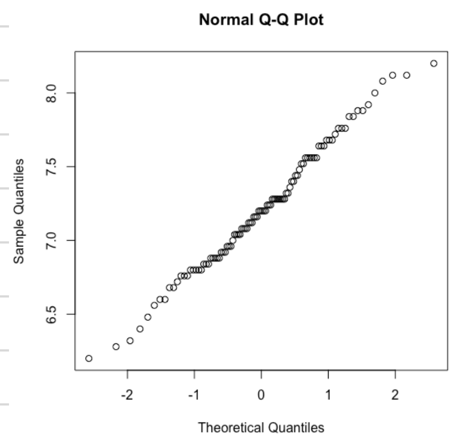




For n=50:

mean(temp1) = 0.99607   mean(temp2) = 7.2104.





d). We can see for $n=20$ and 50.

the mean of both Random vector are closer to the true mean.

Also, the line is "closer" to the normal line with little shift. and less outlier.

which corresponds to the statement in CLT:

As $n \to \infty$, the sample mean converge almost surely to expected value μ.