STA 532 Homework5

Jae Hyun Lee, jl914

17 February, 2020

HW5 for STA-532

1.

(a)

Obtain the MGFs for the chi-square, exponential, and Gamma distributions.

Let X ~ chi-square(k), Y ~ exponential(θ), Z ~ Gamma(a,b).

Then their pdfs are

$$
\begin{aligned}
P_X(x) &= \frac{1}{\Gamma(k/2)} 2^{-k/2} x^{k/2-1} e^{-x/2} \quad for\ x \geq 0 \\
P_y(y) &= \theta e^{-\theta y} \\
P_Z(z) &= \frac{1}{\Gamma(a)} b^a z^{a-1} e^{-bx}
\end{aligned}
$$

and their MGFs are

$$
\begin{aligned}
M_X(t) = E(e^{tx}) &= \int \frac{1}{\Gamma(k/2)} 2^{-k/2} x^{k/2-1} e^{-x/2+tx} dx \\
&= \frac{1}{\Gamma(k/2)} 2^{-k/2} \int \underbrace{x^{k/2-1} e^{-x(1/2-t)}}_{kernel\ of\ gamma(k/2,1/2-t)} dx \\
&= \frac{1}{\Gamma(k/2)} 2^{-k/2} \times \Gamma(k/2) \times (1/2-t)^{-k/2} \\
&= (1-2t)^{-k/2} \\
M_Y(t) = \int e^{ty} \theta e^{-\theta y} dy &= \int \theta e^{-y(\theta - t)} dy = \frac{\theta}{\theta - t} = (1-t/\theta)^{-1} \\
M_Z(t) = E(e^{tz}) &= \int \frac{1}{\Gamma(a)} b^a \underbrace{z^{a-1} e^{-z(b-t)}}_{kernel\ of\ gamma(a,b-t)} dz \\
&= \frac{1}{\Gamma(a)} b^a \times \Gamma(a) \times (b-t)^{-a} \\
&= (1-t/b)^{-a}
\end{aligned}
$$

(b)

Find the distribution of $\sum X_i = \mathbf{X}, \sum Y_i = \mathbf{Y}$

$$
\begin{aligned}
M_{\mathbf{X}}(t) &= E(e^{t\mathbf{X}}) = E(e^{t\sum X_i}) = \prod_{i=1}^{n} E(e^{tX_i}) = (1-2t)^{-nk/2} \sim \chi^2(nk) \\
M_{\mathbf{Y}}(t) &= E(e^{t\mathbf{Y}}) = E(e^{t\sum Y_i}) = \prod_{i=1}^{n} E(e^{tY_i}) = (1-t/\theta)^{-n} \sim Gamma(n, \theta)
\end{aligned}
$$

(c)

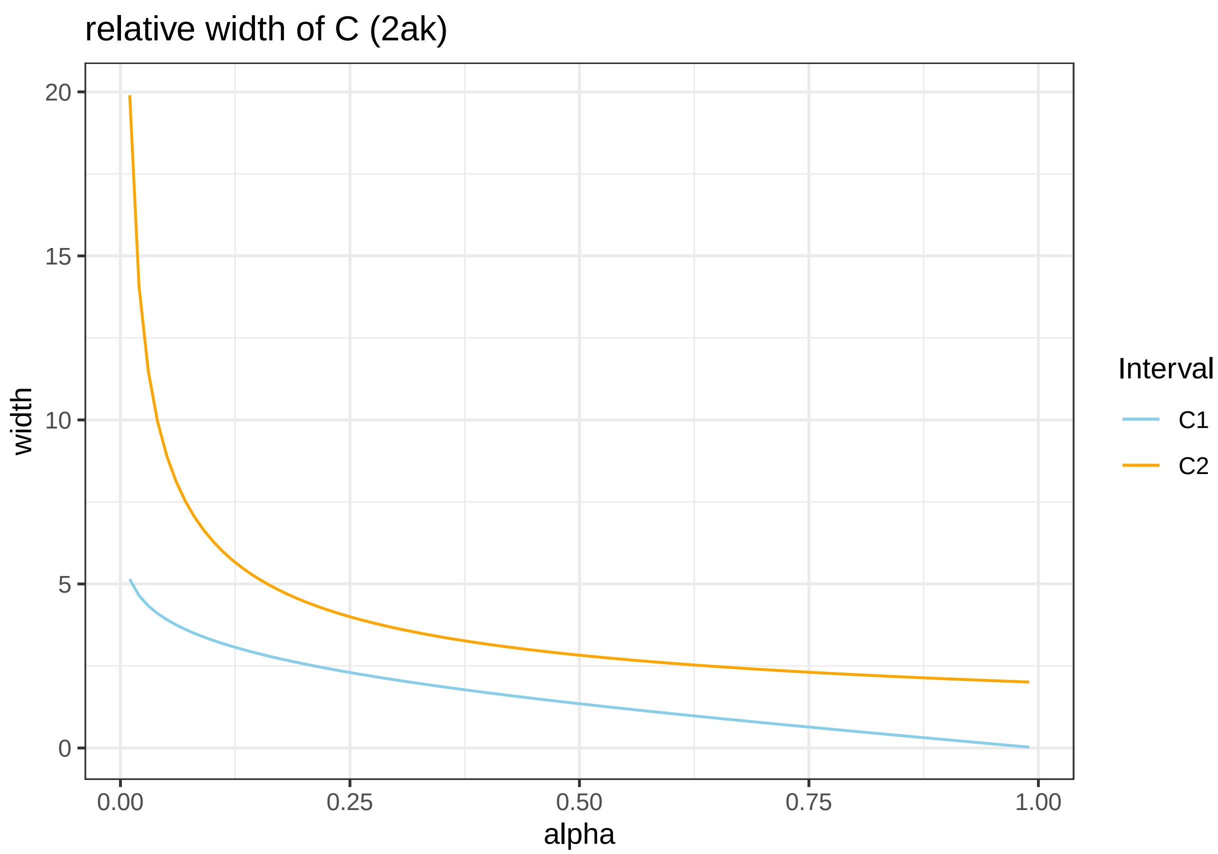
Based on the MGFs, how are the χ^2 and Gamma distributions related?

$$M_X(t) = (1-2t)^{-k/2}$$
$$M_Z(t) = (1-t/\theta)^{-a}$$

It means that χ^2 is special case of Gamma that a = k/2, b = 1/2

2.

```
alpha <- seq(0,1,length.out = 100); alpha <- alpha[-c(1,length(alpha))]
a_norm <- qnorm(p = 1-alpha/2)
a_chevy <- 1/sqrt(alpha)
ggplot() +
  geom_line(mapping = aes(x = alpha, y = 2*a_norm, color = "C1")) +
  geom_line(mapping = aes(x = alpha, y = 2*a_chevy, color = "C2")) +
  theme_bw() +
  labs(title = "relative width of C (2ak)", y = "width", x = "alpha") +
  scale_color_manual("Interval",breaks = c("C1","C2"),values = c("skyblue","orange"))
```



(a)

```r
set.seed(100)
a1 <- qnorm(1-0.8/2)
a2 <- 1/sqrt(0.8)
CI_1 <- matrix(rep(NA,2000),ncol = 2)
coverage_1 <- rep(NA,10)
CI_2 <- matrix(rep(NA,2000),ncol = 2)
coverage_2 <- rep(NA,10)
for(i in 1:10){
  for(j in 1:1000){
    y <- rnorm(i,0,1)
    CI_1[j,] <- c(mean(y) - a1*1/sqrt(i), mean(y) + a1*1/sqrt(i))
    CI_2[j,] <- c(mean(y) - a2*1/sqrt(i), mean(y) + a2*1/sqrt(i))
  }
  coverage_1[i] <- mean(apply(CI_1,1, function(x){ifelse(x[1]<0 & x[2]>0,1,0)}))
  coverage_2[i] <- mean(apply(CI_2,1, function(x){ifelse(x[1]<0 & x[2]>0,1,0)}))
}
kable(rbind(coverage_1,coverage_2), caption = "Y sample from N(0,1)", col.names = paste("n=",1:10))
```

Table 1: Y sample from N(0,1)

	n= 1	n= 2	n= 3	n= 4	n= 5	n= 6	n= 7	n= 8	n= 9	n= 10
coverage_1	0.185	0.195	0.224	0.227	0.200	0.192	0.181	0.199	0.188	0.190
coverage_2	0.731	0.717	0.767	0.762	0.756	0.760	0.699	0.727	0.727	0.722

We can find that coverage of norm, is about 1 - α = 0.2. On the other hand, coverage of chebyshev interval, coverage is much larger than 1 - α = 0.2

(b)

```r
set.seed(100)
library(rmutil)
CI_1 <- matrix(rep(NA,2000),ncol = 2)
coverage_1 <- rep(NA,10)
CI_2 <- matrix(rep(NA,2000),ncol = 2)
coverage_2 <- rep(NA,10)
for(i in 1:10){
  for(j in 1:1000){
    y <- rlaplace(i,0,1)
    CI_1[j,] <- c(mean(y) - a1*sqrt(2)/sqrt(i), mean(y) + a1*sqrt(2)/sqrt(i))
    CI_2[j,] <- c(mean(y) - a2*sqrt(2)/sqrt(i), mean(y) + a2*sqrt(2)/sqrt(i))
  }
  coverage_1[i] <- mean(apply(CI_1,1, function(x){ifelse(x[1]<0 & x[2]>0,1,0)}))
  coverage_2[i] <- mean(apply(CI_2,1, function(x){ifelse(x[1]<0 & x[2]>0,1,0)}))
}
kable(rbind(coverage_1,coverage_2), caption = "Y sample from Laplace(0,1)", col.names = paste("n=",1:10)
```

Table 2: Y sample from Laplace(0,1)

	n= 1	n= 2	n= 3	n= 4	n= 5	n= 6	n= 7	n= 8	n= 9	n= 10
coverage_1	0.301	0.227	0.209	0.218	0.220	0.202	0.191	0.203	0.218	0.207

	n= 1	n= 2	n= 3	n= 4	n= 5	n= 6	n= 7	n= 8	n= 9	n= 10
coverage_2	0.792	0.778	0.771	0.773	0.762	0.773	0.744	0.743	0.765	0.728

Similar with previous result, we could find that coverage based on normal interval is about 0.2 which is 1 - α. On contrary, coverage of chebyshev interval is much larger than 0.2.

(c)

```
set.seed(100)
CI_1 <- matrix(rep(NA,2000),ncol = 2)
coverage_1 <- rep(NA,10)
CI_2 <- matrix(rep(NA,2000),ncol = 2)
coverage_2 <- rep(NA,10)
beta_var <- .1*.5/((.1+.5)^2*(.1+.5+1))
for(i in 1:10){
  for(j in 1:1000){
    y <- rbeta(i,.1,.5)
    CI_1[j,] <- c(mean(y) - a1*sqrt(beta_var)/sqrt(i), mean(y) + a1*sqrt(beta_var)/sqrt(i))
    CI_2[j,] <- c(mean(y) - a2*sqrt(beta_var)/sqrt(i), mean(y) + a2*sqrt(beta_var)/sqrt(i))
  }
  coverage_1[i] <- mean(apply(CI_1,1, function(x){ifelse(x[1]<0 & x[2]>0,1,0)}))
  coverage_2[i] <- mean(apply(CI_2,1, function(x){ifelse(x[1]<0 & x[2]>0,1,0)}))
}
kable(rbind(coverage_1,coverage_2), caption = "Y sample from beta(0.1,0.5)", col.names = paste("n=",1:1
```

Table 3: Y sample from beta(0.1,0.5)

	n= 1	n= 2	n= 3	n= 4	n= 5	n= 6	n= 7	n= 8	n= 9	n= 10
coverage_1	0.680	0.506	0.372	0.261	0.199	0.145	0.129	0.079	0.044	0.042
coverage_2	0.805	0.669	0.618	0.530	0.497	0.414	0.403	0.325	0.290	0.272

For both interval, coverages are dramatically decreased as n increase. However, coverage of chebyshev interval still above $0.2 = 1 - \alpha$ at n= 10. However, coverage of normal interval is much lower than 0.2 when n is larger than 5.

By observing above results, we can infer that z-interval is not robust as much as chebyshev interval because it seems to work well only the case when the distribution is similar with normal distribution like Laplace distribution. On the other hand, we could find that the robustness of chebyshev interval which keeps showing larger coverage than we expect 1-α.

3.

(a)

Pdf and MGF of Y are as follow:

$$
\begin{aligned}
&P(Y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{Y^2}{2\sigma^2}} \\
\rightarrow &M_Y(t) = \int \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(Y^2 - 2\sigma^2 tY + \sigma^4 t^2)} e^{\frac{\sigma^4 t^2}{2}} = e^{\frac{\sigma^4 t^2}{2}} \\
&M_Y'(t) = (\sigma^2 t) e^{\frac{\sigma^2 t^2}{2}}, M_Y''(t) = \sigma^2 e^{\frac{\sigma^2 t^2}{2}} + (\sigma^4 t^2) e^{\frac{\sigma^2 t^2}{2}} \\
&M_Y^{(3)}(t) = \sigma^4 t e^{\frac{\sigma^2 t^2}{2}} + 2\sigma^4 t e^{\frac{\sigma^2 t^2}{2}} + (\sigma^6 t^3) e^{\frac{\sigma^2 t^2}{2}} \\
&M_Y^{(4)}(t) = \sigma^4 e^{\frac{\sigma^2 t^2}{2}} + \sigma^6 t^2 e^{\frac{\sigma^2 t^2}{2}} + 2\sigma^4 e^{\frac{\sigma^2 t^2}{2}} + 2\sigma^6 t^2 e^{\frac{\sigma^2 t^2}{2}} + 3\sigma^6 t^2 e^{\frac{\sigma^2 t^2}{2}} + \sigma^8 t^4 e^{\frac{\sigma^2 t^2}{2}} \\
&M_Y''(t) = \sigma^2 \rightarrow E(Y_i^2) = \sigma^2 \\
&M_Y^{(4)}(0) = 3\sigma^4 \rightarrow Var(Y_i) = E(Y_i^4) - E(Y_i^2) = 2\sigma^4
\end{aligned}
$$

$$
\begin{aligned}
E(\bar{Y}^2) &= E(\sum Y_i^2)/n = \frac{1}{n}\sum E(Y_i^2) = n/n\sigma^2 = \sigma^2 \\
Var(\bar{Y}^2) &= \frac{1}{n^2}\sum Var(Y_i) = \frac{1}{n^2} 2\sigma^4 \times n = \frac{2}{n}\sigma^4
\end{aligned}
$$

(b)

$$
\begin{aligned}
&Pr(\mid \bar{Y}^2 - E(\bar{Y}^2) \mid > \epsilon) \leq Var(\bar{Y}^2)/\epsilon^2 \quad by\ Chebyshev\ inequality. \\
\rightarrow &Pr(\mid \bar{Y}^2 - \sigma^2 \mid > \epsilon) \leq \frac{2\sigma^4}{n\epsilon^2} \rightarrow 0 \quad as\ n \rightarrow \infty \\
\rightarrow &\bar{Y}^2 \rightarrow^p \sigma^2
\end{aligned}
$$

$\bar{Y}^2$ converges in probability to σ^2

(c)

In above questions, we have confirmed that

$$
\begin{aligned}
&Y_1^2, Y_2^2 \cdots Y_n^2 \sim iid\ P, \quad E(\bar{Y}^2) = \sigma^2, \quad V(\bar{Y}^2) = \frac{2}{n}\sigma^4. \\
&As\ n \rightarrow \infty, by\ CLT \\
&\bar{Y}^2 \sim N(\sigma^2, \frac{2}{n}\sigma^4) \quad with\ mean\ \sigma^2 \quad and\ variance\ \frac{2}{n}\sigma^4
\end{aligned}
$$

Exercise 4.

(a)

$$
\begin{aligned}
E(\bar{Y_w}) &= E(\sum w_i Y_i) = \sum w_i E(Y_i) = \mu \sum w_i = \mu \\
Var(\bar{Y_w}) &= Var(\sum w_i Y_i) = \sum V(w_i Y_i) = \sum w_i^2 Var(Y_i) = \sigma^2 \sum w_i^2 a_i
\end{aligned}
$$

(b)

We need to find w_i's minimize $\sigma^2 \sum w_i^2 a_i$. Let $v_i = w_i\sqrt{a_i}$ and $\mathbf{v} = (v_1, v_2, \cdots, v_n)$, $1_n = (1/\sqrt{n}, \cdots, 1/\sqrt{n})$. Then by Cauchy-Swartz inequality,

$$\mathbf{v} \cdot 1_n \leq || \mathbf{v} || \times || 1_n || \rightarrow (\mathbf{v} \cdot 1_n)^2 \leq \sum w_i^2 a_i$$

and equality holds when $\mathbf{v}$ is multiple of 1_n, that means

$$w_i \sqrt{a_i} = k/\sqrt{n} \rightarrow w_i = k/\sqrt{na_i} \quad and \quad we\ know\ that$$
$$\sum w_i = 1 \rightarrow \sum k/\sqrt{na_i} = 1 \rightarrow k = (\sum 1/\sqrt{na_i})^{-1}$$
$$which\ converges\ to\ 0\ as\ n \rightarrow 0$$
$$\rightarrow w_i = (\frac{1}{\sqrt{na_i}})(\sum \frac{1}{\sqrt{na_i}})^{-1}$$

(c)

As we have discussed previous question k converges to 0 as n increase. Thus

$$Pr(| \bar{Y_w} - E(\bar{Y_w}) |> \epsilon) \leq \frac{Var(\bar{Y_w})}{\epsilon^2} \quad (by\ chebyshev)$$
$$\frac{Var(\bar{Y_w})}{\epsilon^2} = \sigma^2/\epsilon^2 \times \sum w_i^2 a_i = \frac{\sigma^2 k^2}{n\epsilon^2} \rightarrow 0 \quad as\ n \rightarrow \infty$$

Thus this is WLLN for $\bar{Y_w}$.

Exercise 5

(a)

Let Pr(Y < y) = F(y) = p where $0 \leq p \leq 1$. $\rightarrow F_{Y_i}(y) = Pr(Y_i \leq y)$. Then $\hat{F(y)} = \hat{p} = \frac{1}{n}\sum Z_i$ where $Z_i = 1$ when $Y_i < y$ with probability p and otherwise $Z_i = 0$. This indicates that $Z_i \sim binary(p)$. Then $E(Z_i) = p, Var(Z_i) = p(1-p)$

$$E(\hat{F(y)}) = E(\hat{p}) = \frac{1}{n}\sum E(Z_i) = \frac{1}{n}np = p = F(y) \rightarrow \hat{F(y)}\ is\ unbiased\ estimator\ of\ F(y)$$

Moreover,

$$Pr(| \hat{p} - p |> \epsilon) \leq Var(\hat{p})/\epsilon^2 = p(1-p)/n\epsilon^2 \rightarrow 0 \quad as\ n \rightarrow \infty$$

which means that $\hat{p} \rightarrow^p p$ and it means that $\hat{p}$ is consistent estimator for p.

It variance $Var(\hat{p}) = \frac{1}{n^2}np(1-p) = p(1-p)/n = F(y)/n$

(b)

We have checked that $\hat{p} \rightarrow^p p$ and $g(\hat{p}) \rightarrow^p g(p)$ for g is continous. $g(x) = \sqrt{x(1-x)} \quad where\ 0 < x < 1$ which is contious. Therefore, for large n

$$\frac{\sqrt{n}(\hat{p}-p)}{\sqrt{\hat{p}(1-\hat{p})}} = \frac{\sqrt{n}(\hat{p}-p)}{\sqrt{p(1-p)}} \times \frac{\sqrt{p(1-p)}}{\sqrt{\hat{p}(1-\hat{p})}}$$
$$where\ \frac{\sqrt{n}(\hat{p}-p)}{\sqrt{p(1-p)}} \rightarrow^d N(0,1) \quad by\ CLT, \quad \frac{\sqrt{p(1-p)}}{\sqrt{\hat{p}(1-\hat{p})}} \rightarrow^p 1\ by\ consistency$$
$$Thus\ \frac{\sqrt{n}(\hat{p}-p)}{\sqrt{\hat{p}(1-\hat{p})}} \rightarrow^d N(0,1)$$

CI for normal distribution: we use $Z_{\alpha/2}$.

Therefore, 95% CI of F(y) is as follow

$$
\begin{aligned}
Pr(F(y) \in C(F(y))) &= Pr(-Z_{\alpha/2} < \frac{\sqrt{n}(\hat{p}-p)}{\sqrt{\hat{p}(1-\hat{p})}} < Z_{\alpha/2}) \\
&= Pr(-Z_{\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n} < \hat{p}-p < Z_{\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n}) \\
&= Pr(\hat{p}-Z_{\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n} < p < \hat{p}+Z_{\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n}) \\
\rightarrow C(F(y) &= [-Z_{\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n}, Z_{\alpha/2}\sqrt{\hat{p}(1-\hat{p})/n}]
\end{aligned}
$$