

Quiz #1, Solutions

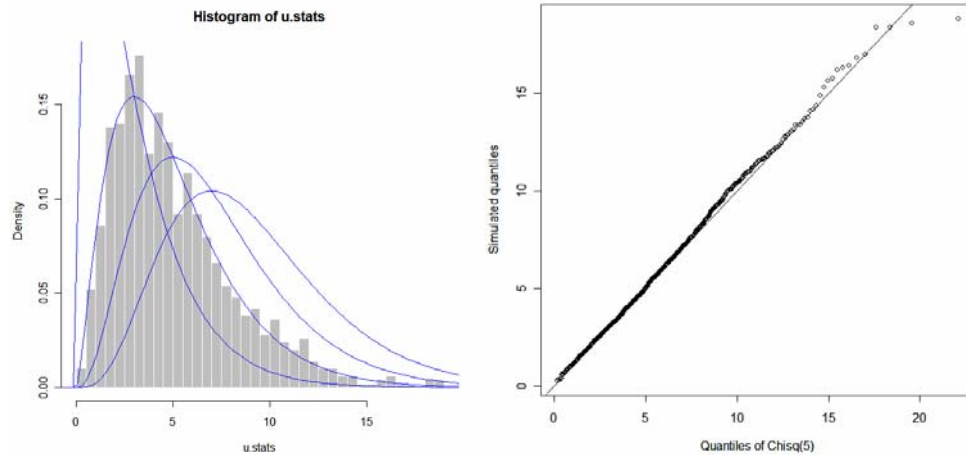
#1.

```
n <- 1000; k <- 5
u.stats <- numeric(n)
for(i in 1:n){
  u.stats[i] <- sum(rnorm(k)^2)
}

hist(u.stats, nclass=50, col='gray', prob=TRUE, border='white')

for(i in seq(3,9,by=2)) curve(dchisq(x, i), add=TRUE, col='blue')

plot(qchisq(ppoints(u.stats),5), sort(u.stats),
     xlab='Quantiles of Chisq(5)', ylab='Simulated quantiles')
abline(0,1)
```



#2.

```
contam <- rnorm(100, 0, (1+9*rbinom(100, 1, 0.05)))
median(contam)
mean(contam)

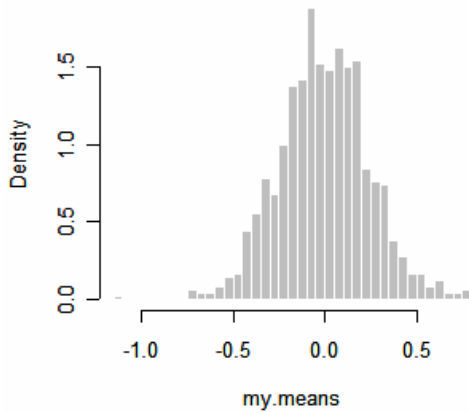
m <- 1000
my.medians <- my.means <- numeric(m)
for(i in 1:m){
  contam <- rnorm(100, 0, (1+9*rbinom(100, 1, 0.05)))
  my.medians[i] <- median(contam)
  my.means[i] <- mean(contam)
}
my.medians[i] <- median(contam)
my.means[i] <- mean(contam)
}

par(mfrow=c(2,2))
hist(my.means, nclass=50, col='gray', prob=TRUE, border='white')
hist(my.medians, nclass=50, col='gray', prob=TRUE, border='white')
qqnorm(my.medians);qqline(my.medians)
qqnorm(my.means);qqline(my.means)

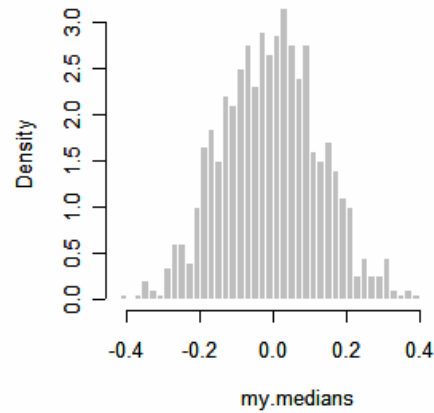
mean(my.medians)
sd(my.medians)
```

```
mean(my.means)
sd(my.means)
```

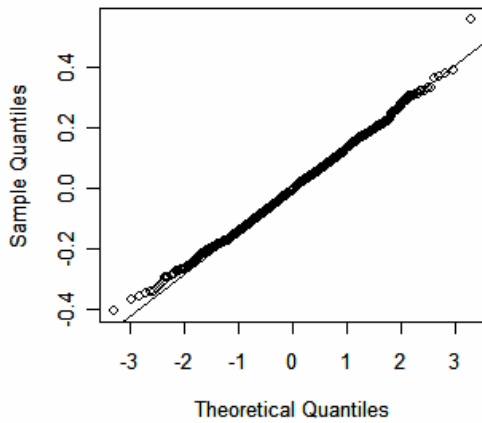
Histogram of my.means



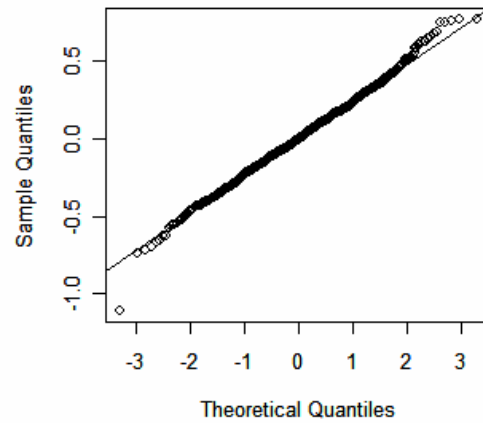
Histogram of my.medians



Normal Q-Q Plot



Normal Q-Q Plot



4) It means their sampling distributions are approximately normal with the given means and SDs. The median is more preferable than the mean for such contaminated normal model since it has smaller SD while their biases are about the same.