Review Problems Unit 0

Advanced Statistical Inference

Alexander J Ohrt

9/30/2021

Setting some of the parameters to arbitrary values.

```
sigma <- 1
mu_1 <- 1
mu_2 <- 2 # To show that both expectations do not have to be equal.
```

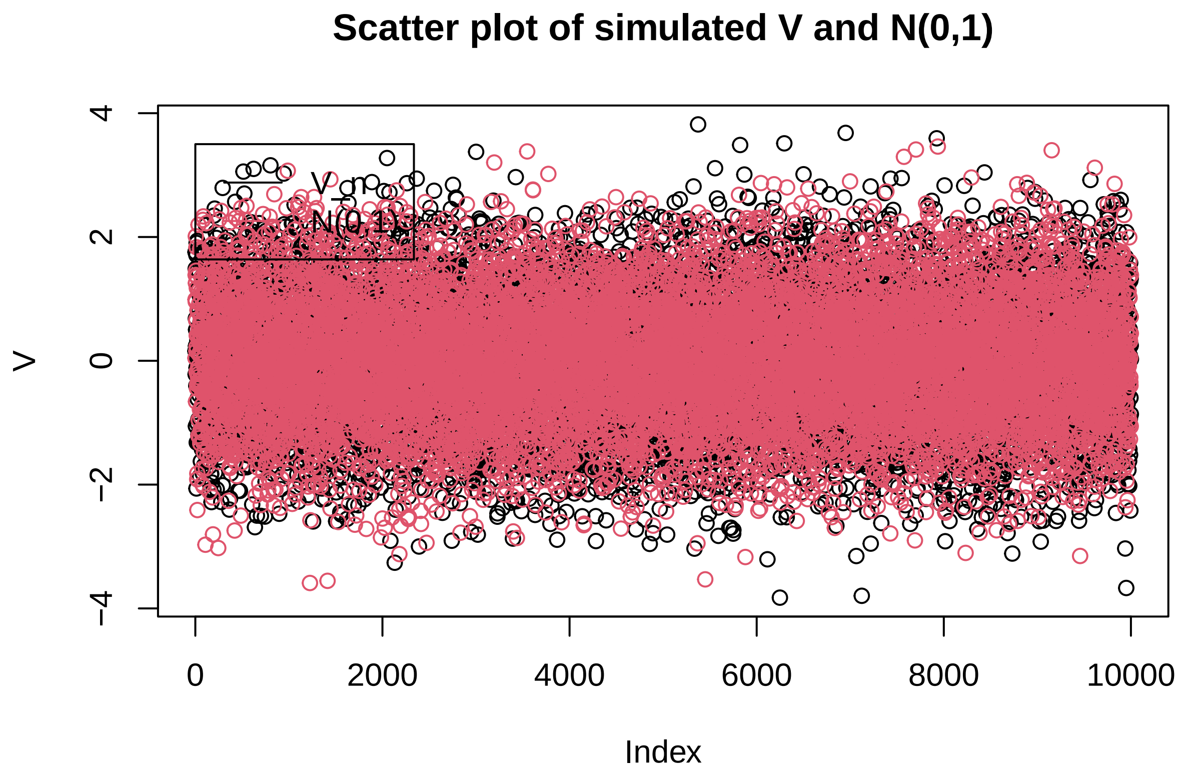
a)

Show that $V_n = \frac{1}{2\sigma}\sqrt{n}(\overline{X} - \overline{Y} - \mu_1 - \mu_2)$ is distributed as $N(0,1)$.

```
N <- 1000
reps <- 10000
V <- rep(0,reps)
for (i in 1:reps){
  # Sampling X and Y from their respective distributions.
  X <- rnorm(N, mu_1, sqrt(sigma^2))
  Y <- rnorm(N, mu_2, sqrt(3*sigma^2))
  V[i] <- sqrt(N)*(mean(X)-mean(Y)-mu_1+mu_2)/(2*sigma)
}
norm <- rnorm(reps)
```

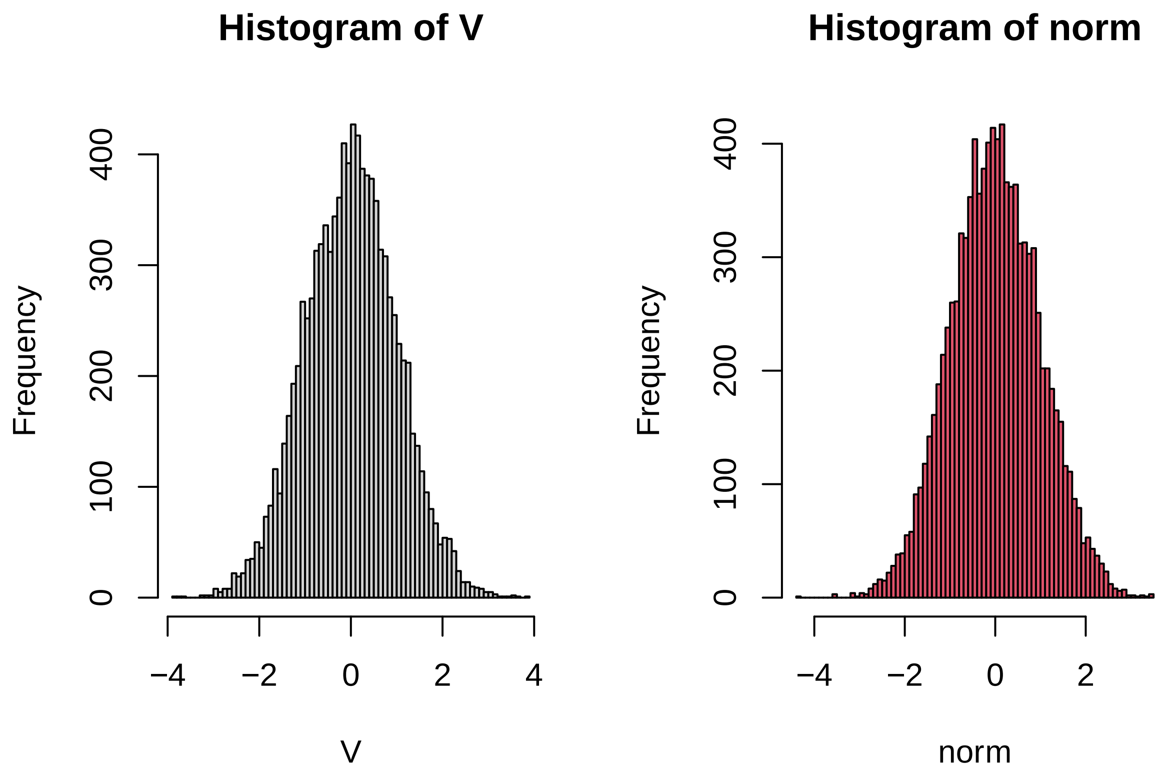
Firstly I make a scatter plot of the points from V and `rnorm(1000)`, to check if there are some obvious differences.

```
par(mfrow=c(1,1))
plot(V, col = 1, lty = 1, main = "Scatter plot of simulated V and N(0,1)")
points(norm, col = 2, lty = 1)
legend(0, 3.5, c("V_n", "N(0,1)"), col = c(1,2), lty = 1)
```



Secondly I make histograms of the two vectors, in order to compare them side by side. The simulation of V_n seems to be distributed (standard) normally, based on the histograms.

```
par(mfrow=c(1,2))
hist(V, breaks = 100)
hist(norm, col = 2, breaks = 100)
```

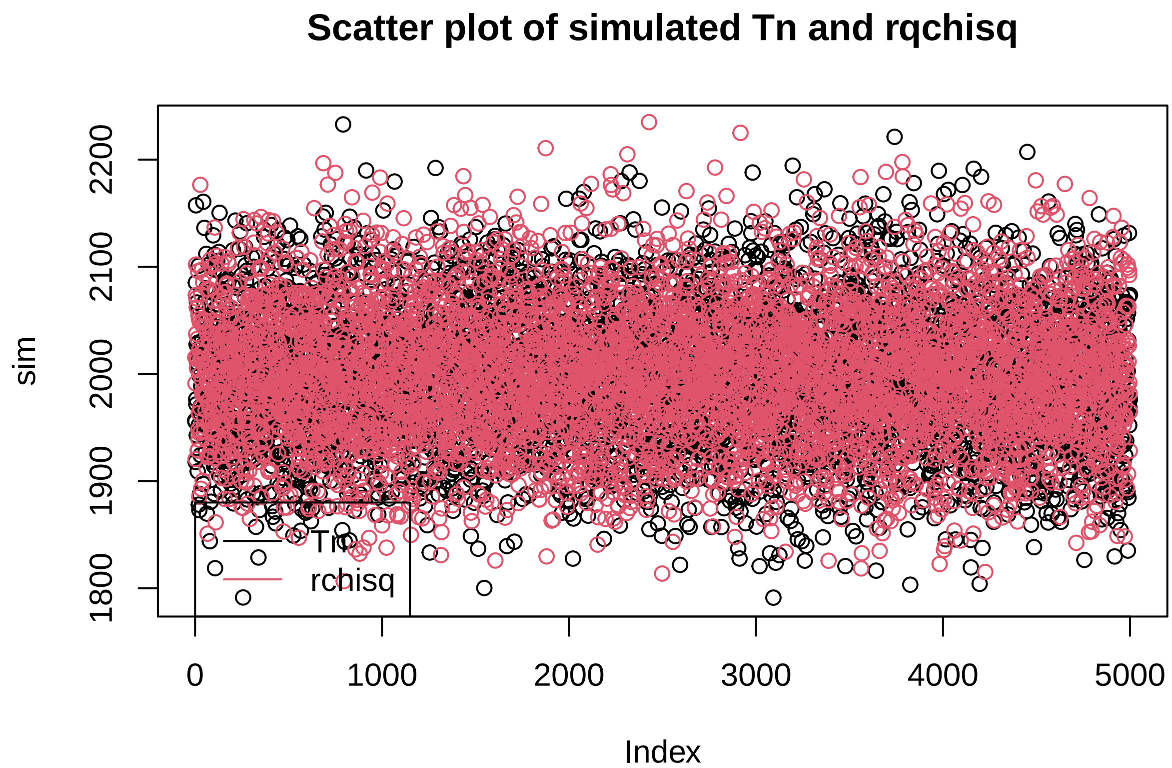


b)

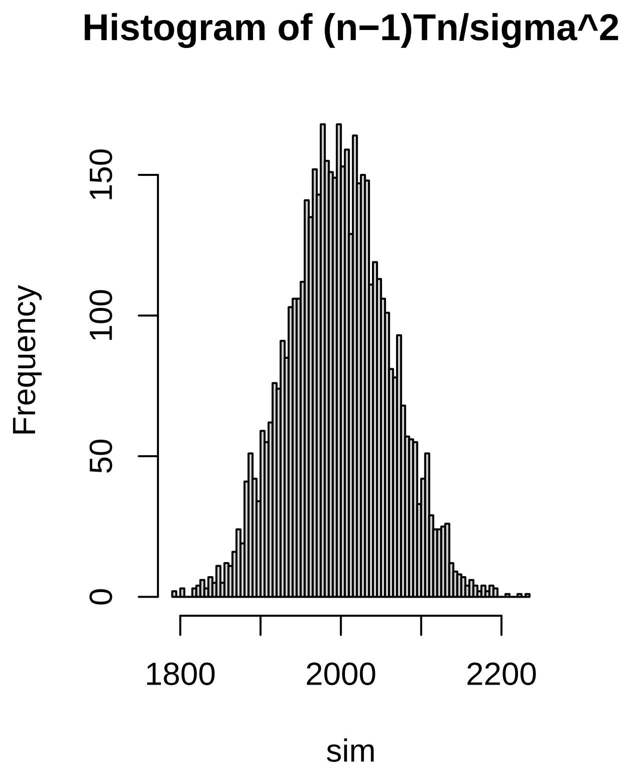
Show that $(n-1)T_n/\sigma^2 = \frac{(n-1)}{\sigma^2}(S_{1n}^2 + \frac{1}{3}S_{2n}^2)$ is distributed as $\chi^2_{2(n-1)}$.

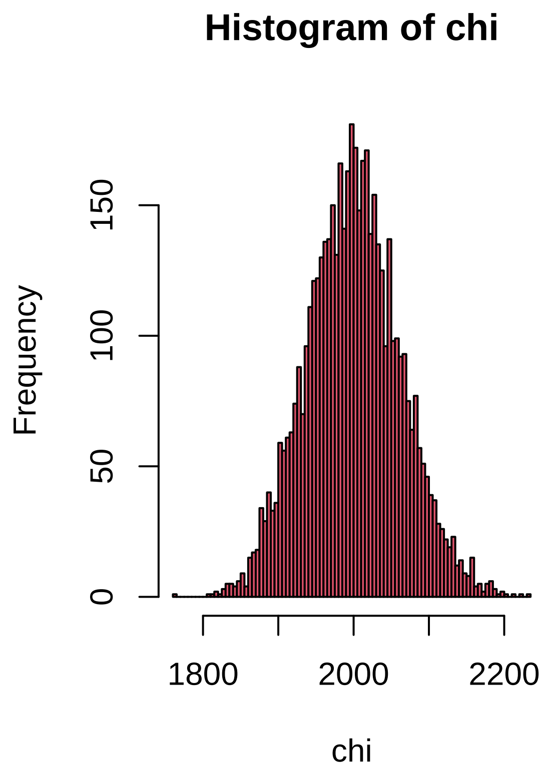
```
reps <- 5000
N <- 1000
sim <- rep(0,reps)
chi <- rep(0,reps)
for (i in 1:reps){
  # Sampling X and Y from their respective distributions.
  X <- rnorm(N, mu_1, sqrt(sigma^2))
  Y <- rnorm(N, mu_2, sqrt(3*sigma^2))
  sim[i] <- (N-1)*(var(X) + 1/3*var(Y))/sigma^2
}
chi <- rchisq(reps, 2*(N-1))
```

```
par(mfrow=c(1,1))
plot(sim, col = 1, lty = 1, main = "Scatter plot of simulated Tn and rqchisq")
points(chi, col = 2, lty = 1)
legend(0, 1880, c("Tn", "rchisq"), col = c(1,2), lty = 1)
```




```
par(mfrow=c(1,2))
hist(sim, breaks = 100, main = "Histogram of (n-1)Tn/sigma^2")
hist(chi, col = 2, breaks = 100)
```





Similar to the conclusion in a), the plots show what we wanted to show.

c)

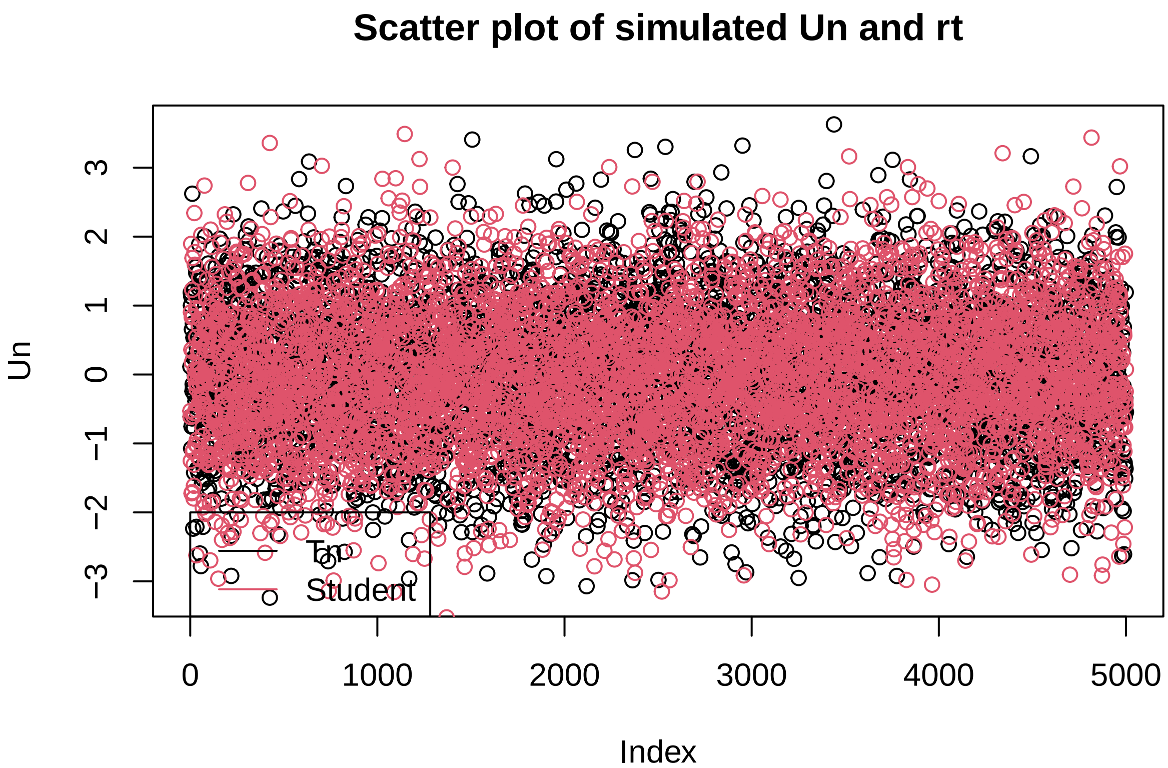
Not sure how to show that two statistics are independent via simulation in R.

d)

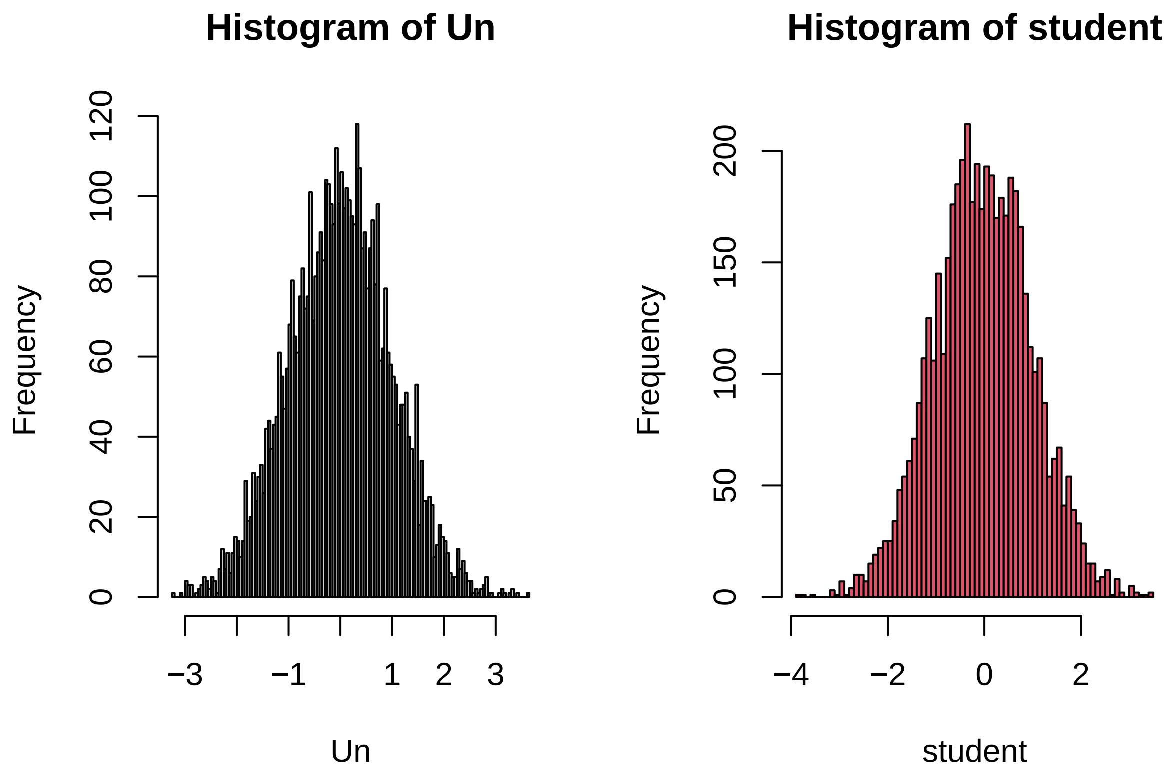
Show that $U_n = \sqrt{n}(\overline{X} - \overline{Y} - \mu_1 + \mu_2)/\sqrt{2T_n} \sim t_{2(n-1)}$.

```r
reps <- 5000
N <- 1000
Tn <- rep(0,reps)
Un <- rep(0,reps)
for (i in 1:reps){
  # Sampling X and Y from their respective distributions.
  X <- rnorm(N, mu_1, sqrt(sigma^2))
  Y <- rnorm(N, mu_2, sqrt(3*sigma^2))
  Tn[i] <- var(X) + 1/3*var(Y)
  Un[i] <- sqrt(N)*(mean(X)-mean(Y)-mu_1+mu_2)/sqrt(2*Tn[i])
}
student <- rt(reps, 2*(N-1))
```

```r
par(mfrow=c(1,1))
plot(Un, col = 1, lty = 1, main = "Scatter plot of simulated Un and rt")
points(student, col = 2, lty = 1)
legend(0, -2, c("Tn", "Student"), col = c(1,2), lty = 1)
```




```
par(mfrow=c(1,2))
hist(Un, breaks = 100, main = "Histogram of Un")
hist(student, col = 2, breaks = 100)
```



As before, the plots simulate the theoretical conclusions.

e) and f)

Is it possible to simulate (e) (convergence in probability) and (f) (convergence in law)?