Laboration 3

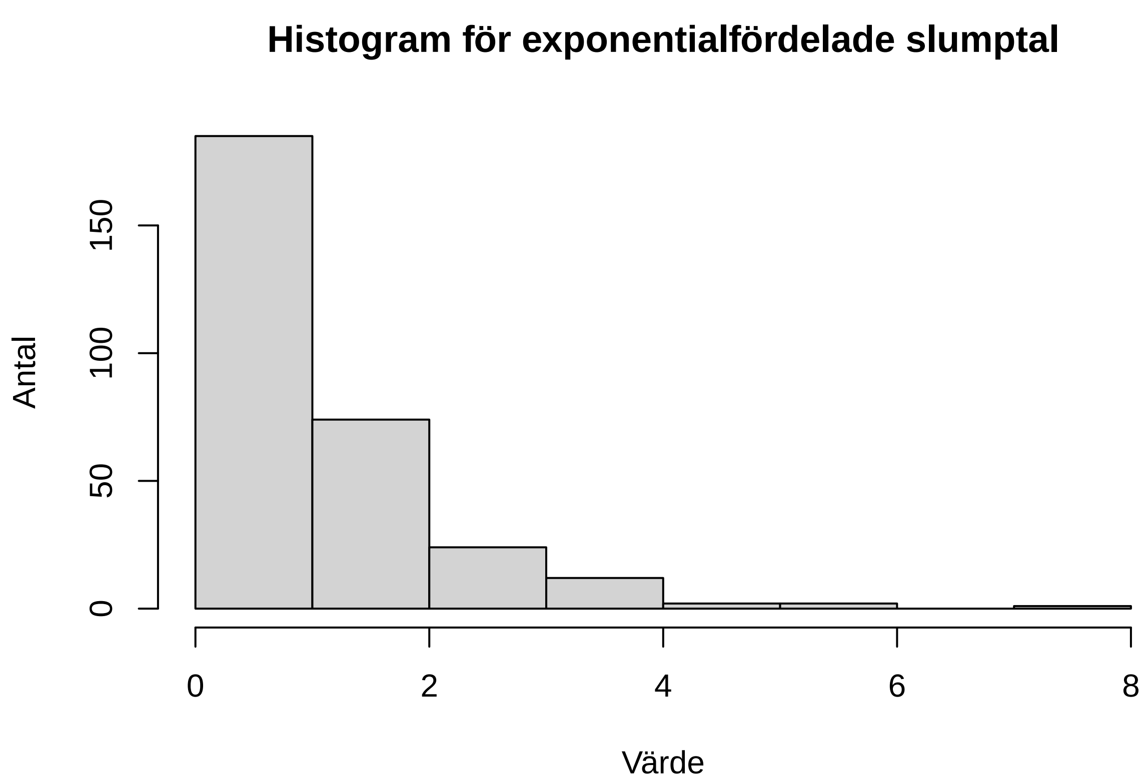
Ville Wassberg

2022-10-28

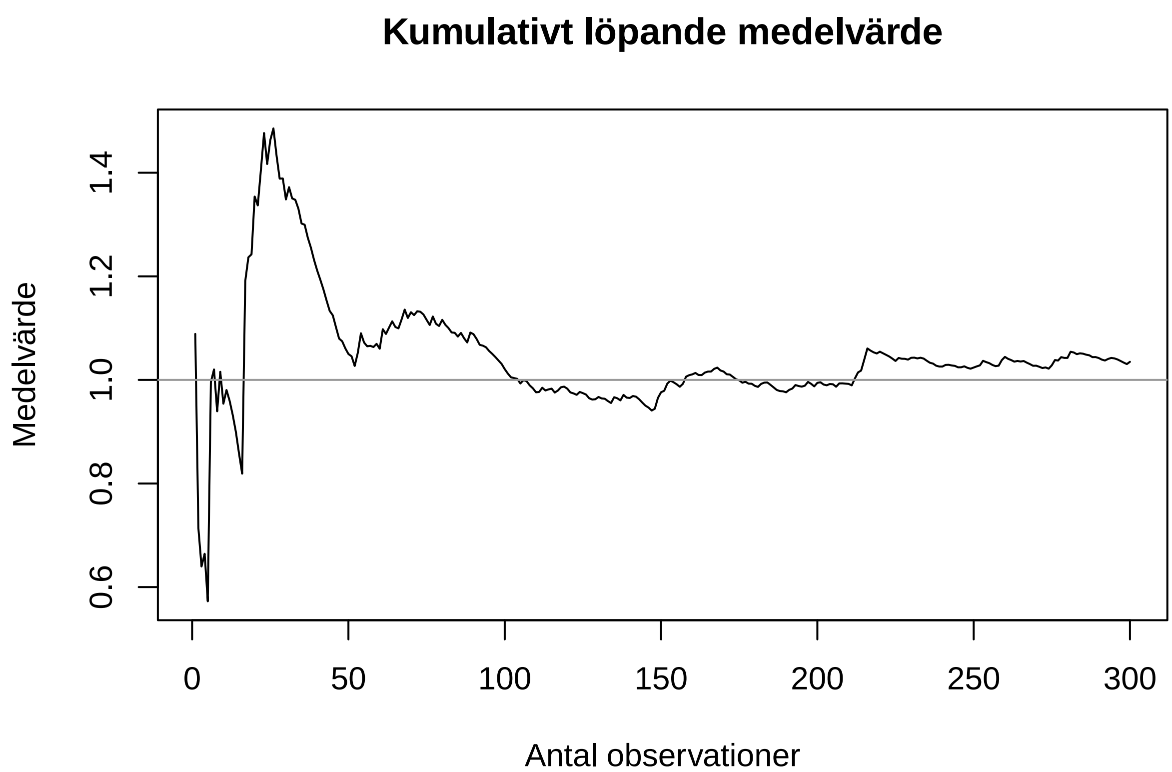
Uppgift 1

De stora talens lag är en sats som beskriver att efter ett stort antal observationer så konvergerar medelvärdet för någon stokastisk variabel mot variabelns väntevärde. Exempelvis om någon ska dra valörer från en kortlek med återläggning så kan det hända att damen dras fler gånger än en på 13, men att så småningom när antalet dragningar är väldigt stort, så¨går det aritmetiska medelvärdet mot en trettondel.

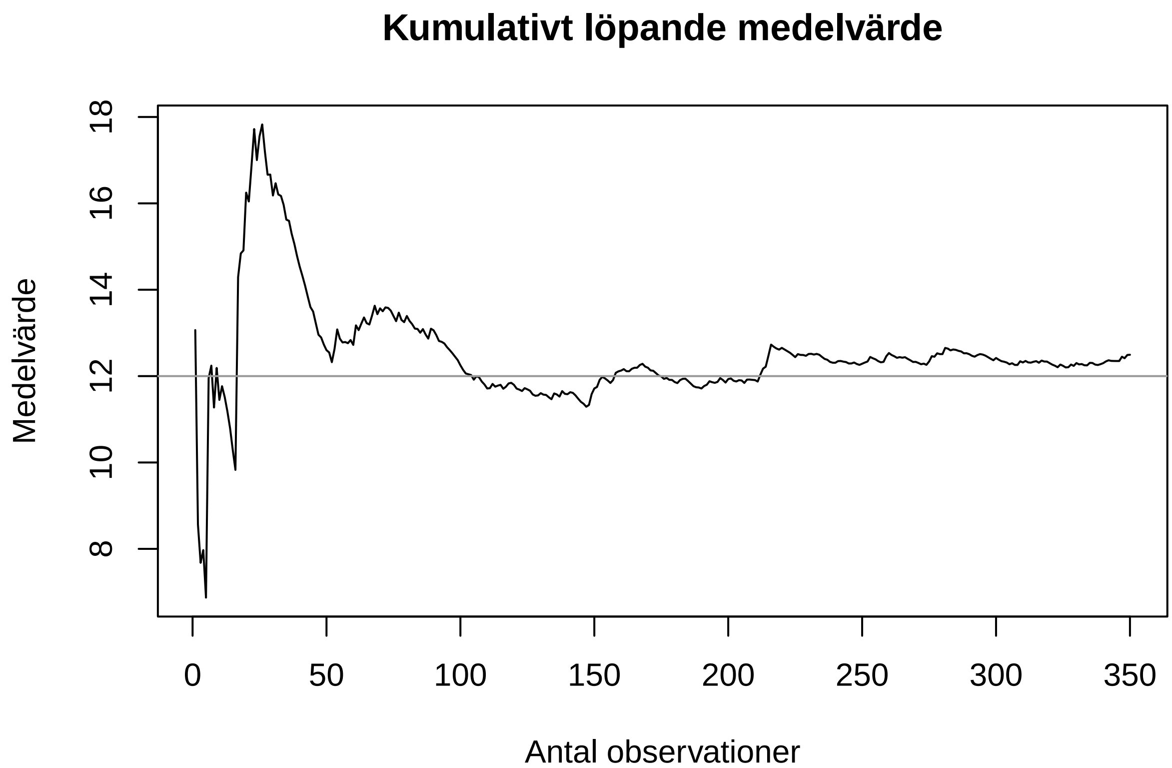
```
set.seed(19930610)
slumptal <- rexp(300, rate = 1)
hist(slumptal,
main = "Histogram för exponentialfördelade slumptal",
ylab = "Antal",
xlab = "Värde")
```



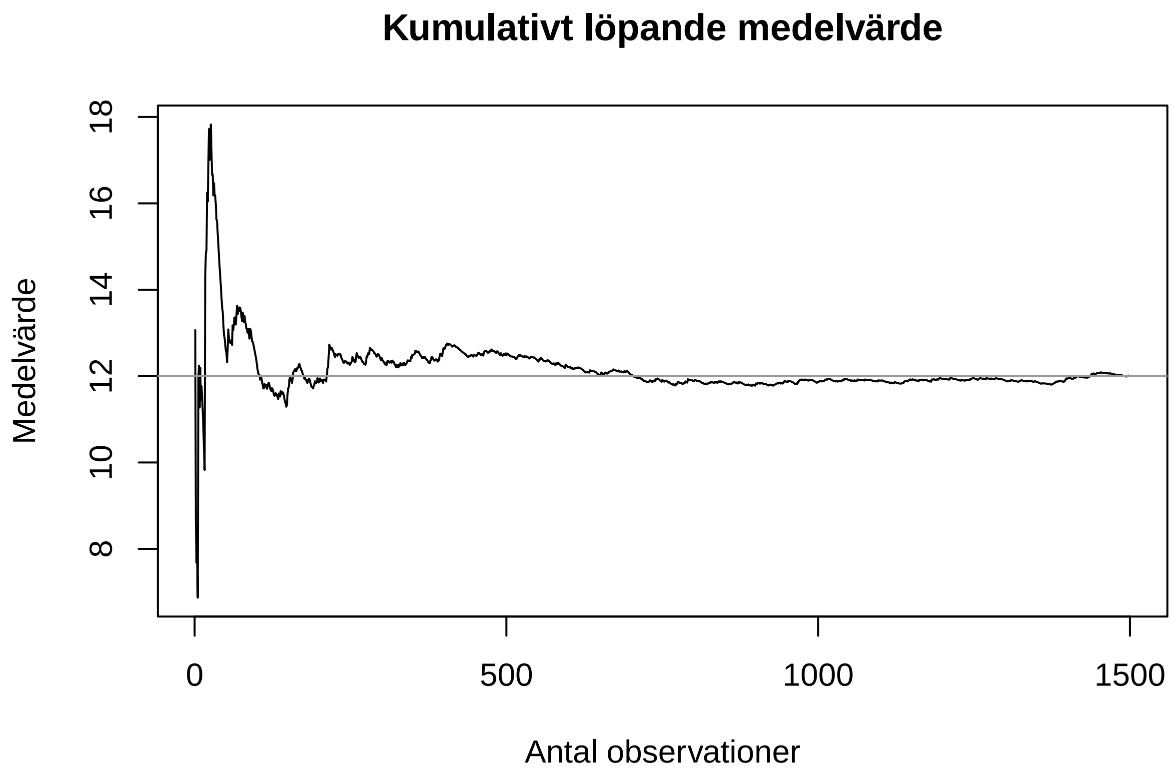
```r
medel <- cumsum(slumptal) / 1:length(slumptal)
plot(medel,
type = "l", # plotta en linje, inte ringar (default)
main = "Kumulativt löpande medelvärde",
ylab = "Medelvärde",
xlab = "Antal observationer")
abline(a = 1, b = 0, col = "gray60") # Horisontell linje vid y = 1 = väntevärdet
```



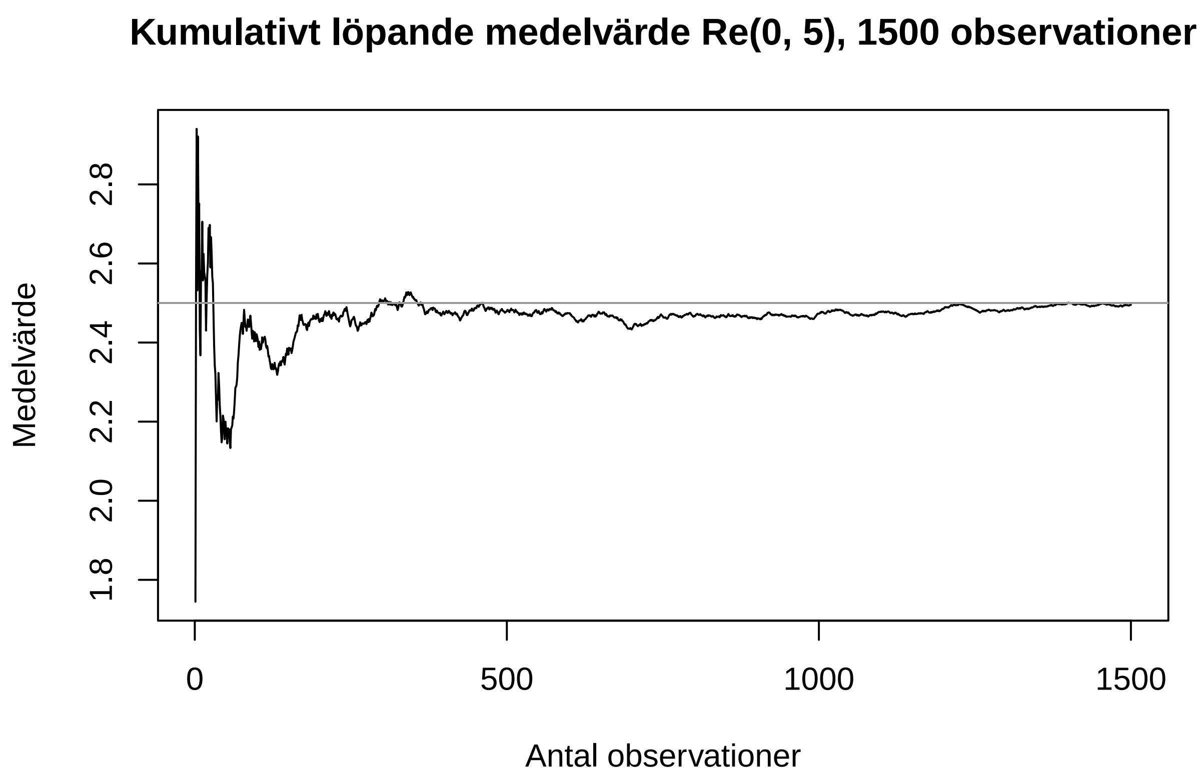

```r
set.seed(19930610)
slumptal2 <- rexp(350, rate = 1/12)
medel2 <- cumsum(slumptal2) / 1:length(slumptal2)
plot(medel2,
type = "l", # plotta en linje, inte ringar (default)
main = "Kumulativt löpande medelvärde",
xlab = "Antal observationer",
ylab = "Medelvärde")
abline(a = 12, b = 0, col = "gray60") # Horisontell linje vid y = 12
```




```
set.seed(19930610)
slumptal3 <- rexp(1500, rate = 1/12)
medel3 <- cumsum(slumptal3) / 1:length(slumptal3)
# Plotta sedan vektorn medel3
plot(medel3,
type = "l", # plotta en linje, inte ringar (default)
main = "Kumulativt löpande medelvärde",
xlab = "Antal observationer",
ylab = "Medelvärde")
abline(a = 12, b = 0, col = "gray60") # Horisontell linje vid y = 12
```




```r
set.seed(19930610)
slumptal4 <- runif(1500, 0, 5) #Likformigt fördelad
medel4 <- cumsum(slumptal4) / 1:length(slumptal4)
plot(medel4,
type = "l",
main = "Kumulativt löpande medelvärde Re(0, 5), 1500 observationer",
xlab = "Antal observationer",
ylab = "Medelvärde")
abline(a = 5/2, b = 0, col = "gray60")
```



Medelvärdet för de exponentialfördelade slumptalen tycks konvergera mot just väntevärdet, med en allt större precision desto högre antal observationer som gjorts. Vilket alltså följer Stora talens lag väldigt väl. Medelvärdet ser allt skakigare ut vid få observationer, vilket var relativt väntat, då, som i exemplet ovan, en händelse kan ske fler gånger än väntat vid få observationer, men vid ett stort antal observationer jämnar det ut sig. Dessa plottar stämmer alltså väldigt väl överens med Stora talens lag.

Även för den likformiga fördelningen konvergerar medelvärdet mot väntevärdet efter ett stort antal observationer, vilket även det var relativt väntat då ett slumptal vid få observationer kan alla ha hamnat i den "övre" halvan av intervallet och öka medelvärdet, men efter många observationer jämnar det ut sig.

Uppgift 2

Centrala gränsvärdessatsen är en sats som säger att summan av n stycken stokastiska variabler med ändlig varians följer normalfördelningskurvan när n är stort nog. Med andra ord, om n går mot oändligheten så följer summan normalfördelningen. Anmärkningsvärt är att vilka fördelningar de stokastiska variablerna följer inte har betydelse.

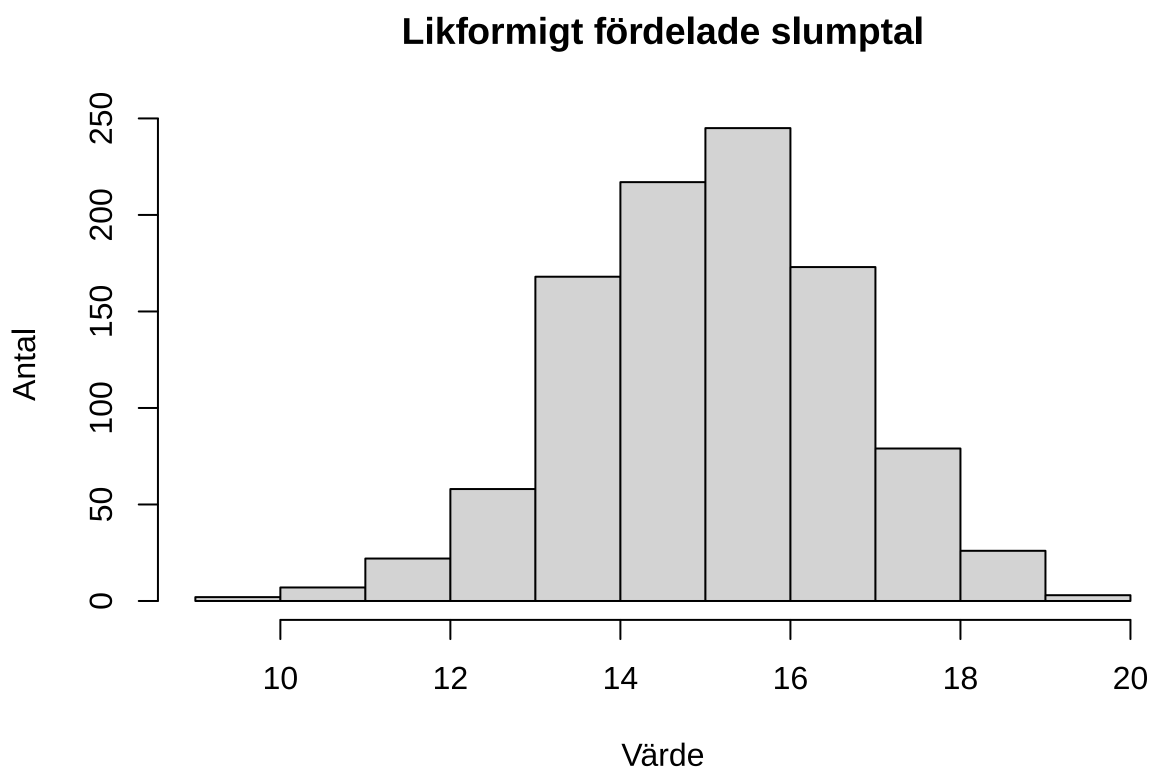
```
set.seed(19930610)
likf <- runif(300)
```

```
set.seed(19930610)
n <- 30 # antalet rader
M <- 1000 # antalet kolumner
likf_matris <- matrix(runif(n * M), nrow = n, ncol = M)
```

```
summa <- colSums(likf_matris)
```

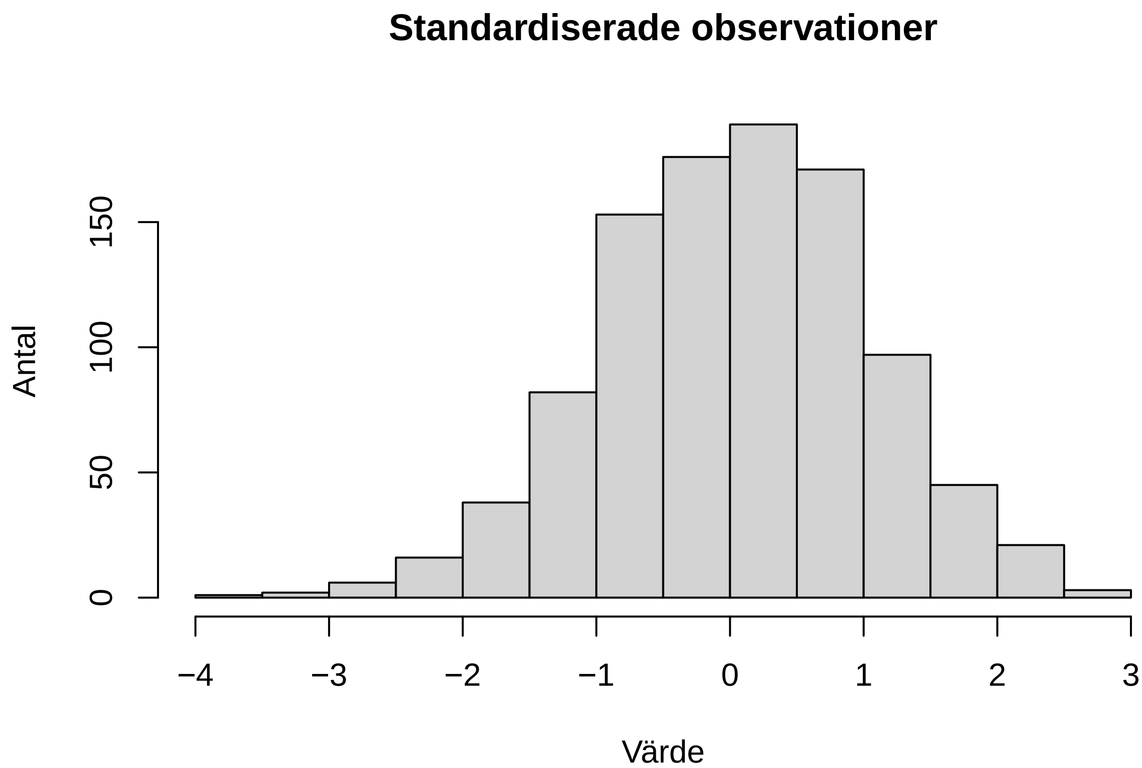
När vektorn summa plottas i ett histogram, så kan observeras att fördelningen ser ut att följa normalfördelningskurvan väldigt väl:

```
hist(summa,
main = "Likformigt fördelade slumptal",
ylab = "Antal",
xlab = "Värde")
```



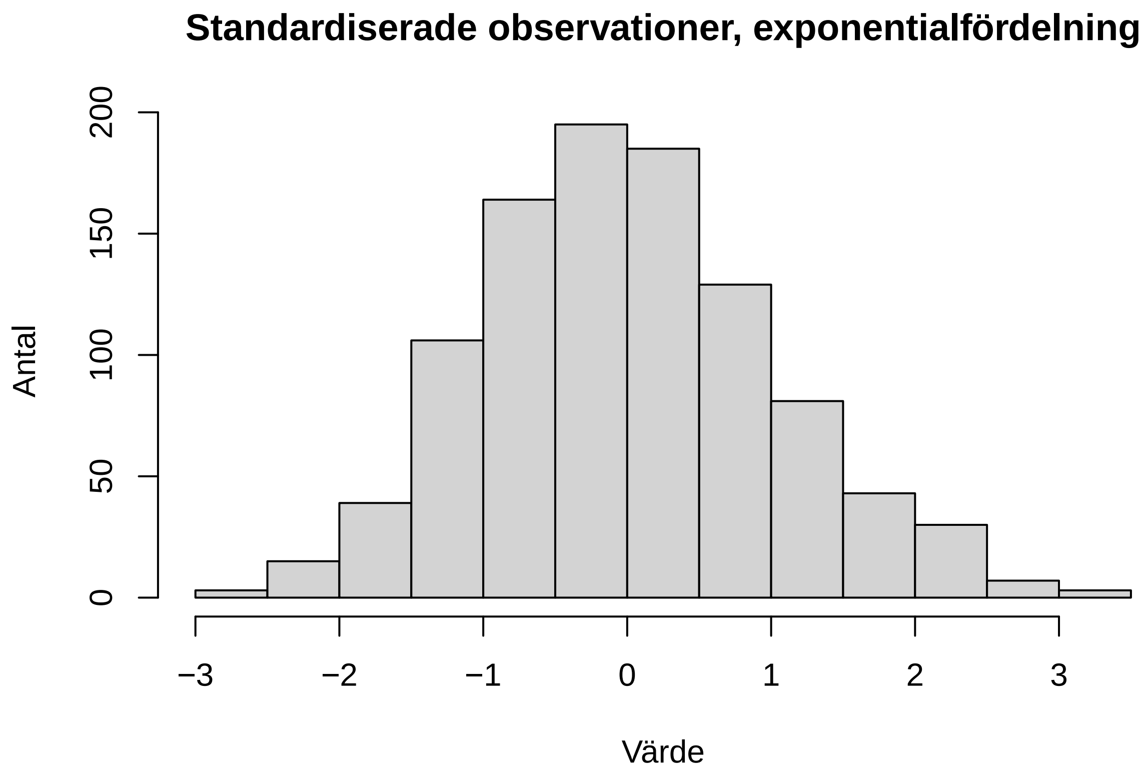
Jämför med normalfördelningskurvan:

```
mu <- 1/2 # Väntevärdet
sigma2 <- 1/12 # Variansen
stand <- (summa - n * mu) / sqrt(n * sigma2)
hist(stand,
main = "Standardiserade observationer",
ylab = "Antal",
xlab = "Värde")
```



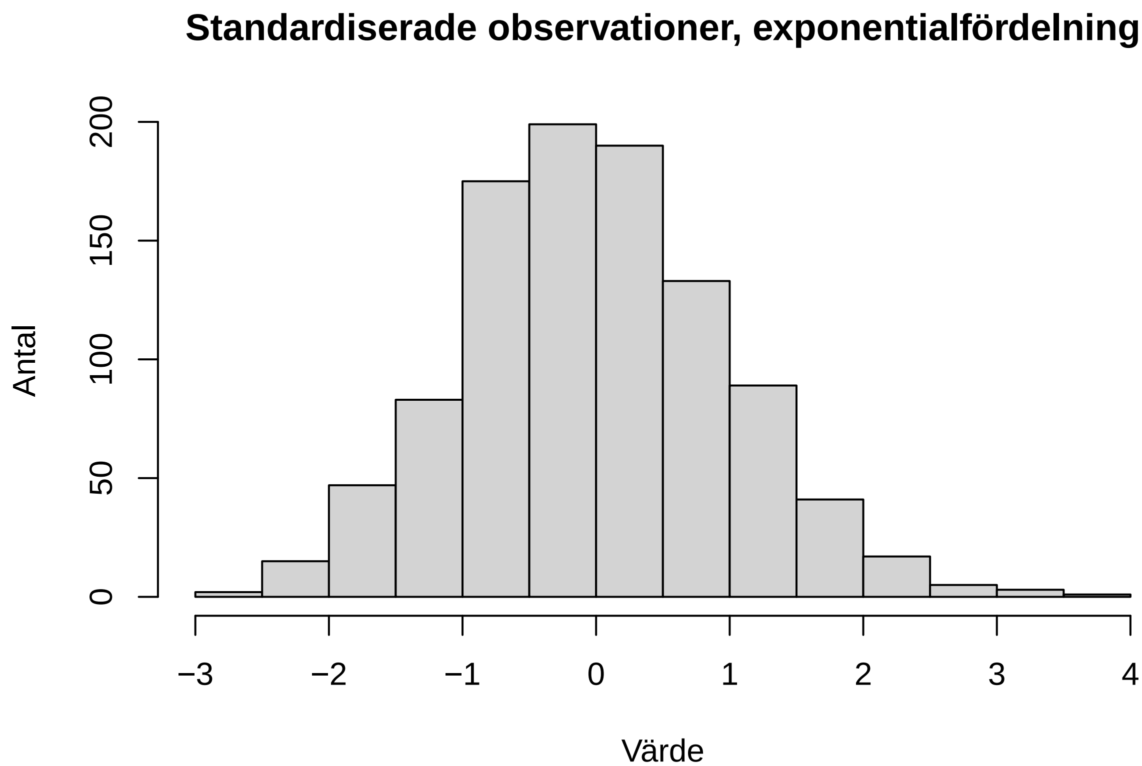
När man sätter n till 30, så kan en se att summan av de stokasiska variablerna tenderar mot normalfördelningskurvan, men att den fortfarande har liknelser med exponentialfördelningen, om än små, där kurvan har pressats in mer centrerat:

```
set.seed(19930610)
n <- 30 # antalet rader
M <- 1000 # antalet kolonner
expo_matris <- matrix(rexp(n * M, rate = 1), nrow = n, ncol = M)
summa <- colSums(expo_matris)
mu <- 1 # Väntevärde, exp
sigma2 <- 1 # Varians, exp
stand <- (summa - n * mu) / sqrt(n * sigma2)
hist(stand,
main = "Standardiserade observationer, exponentialfördelning",
ylab = "Antal",
xlab = "Värde")
```



Med n = 300 blir det ännu lite tydligare, att den, liksom den centrala gränsvärdessatsen säger, tenderar mot normalfördelningskurvan;

```
set.seed(19930610)
n <- 300 # antalet rader
M <- 1000 # antalet kolonner
expo_matris <- matrix(rexp(n * M, rate = 1), nrow = n, ncol = M)
summa <- colSums(expo_matris)
mu <- 1 # Väntevärde, exp
sigma2 <- 1 # Varians, exp
stand <- (summa - n * mu) / sqrt(n * sigma2)
hist(stand,
main = "Standardiserade observationer, exponentialfördelning",
ylab = "Antal",
xlab = "Värde")
```



Slutligen, diagrammen ovan har följt den centrala gränsvärdessatsen väl, då med ett större n har summorna av både likformiga och exponentialfördelade stokastiska variablerna konvergerat mot normalfördelningskurvan. En får dra slutsatsen att plottarna har gått som väntat, då de följer en väletablerad sats.