Exam2

January 9, 2022

1 Load Packages

[100]:
```
using Printf, Statistics, StatsBase, Random, Distributions
include("jlFiles/printmat.jl")
Random.seed!(678)          #set the random number generator to this starting
 ↪point
```

```
TaskLocalRNG()
```

[101]:
```
using Plots

gr(size=(480,320))
default(fmt = :svg)
```

2 Introduction

This exam explores how autocorrelation ought to change how we test statistical hypotheses.

[102]:
```
function SimAR1(T, , , )
    y = fill(NaN, T + 100)
    e = rand(Normal(0, 1), T + 100)
    y[1] = 1
    for i = 2:T + 100
        y[i] = (1- )  +  *y[i-1] + e[i]*
    end
    return y[101:end]
end
```

```
SimAR1 (generic function with 1 method)
```

[103]:
```
y1 = SimAR1(500,0,3,2)

printmat("average from one sample with  =0", mean(y1))

printmat("autocorrelations with  =0")
printmat(1:5, autocor(y1)[2:6])
```

```
average from one sample with  =0     1.986

autocorrelations with  =0

     1        -0.016
     2         0.027
     3         0.004
     4         0.004
     5         0.008
```

[104]:
```
y2 = SimAR1(500,0.75,3,2)

printmat("average from one sample with  =0.75", mean(y2))

printmat("autocorrelations with  =0.75")
printmat(1:5, autocor(y2)[2:6])
```

```
average from one sample with  =0.75     3.187

autocorrelations with  =0.75

     1         0.766
     2         0.580
     3         0.495
     4         0.441
     5         0.372
```

3 Task 2

Do a Monte Carlo simulation. Use the parameters `(T, , , ) = (500,0,3,2)`.

1. Generate a sample with T observations and calculate the average. Repeat $M = 10,000$ times and store the estimated averages in a vector of length M. (The rest of the question uses the symbol μ_i to denote the average from sample i.)
2. What is average μ_i across the M estimates? (That is, what is $\frac{1}{M}\sum_{i=1}^{M}\mu_i$?) *Report* the result.
3. What is the standard deviation of μ_i across the M estimates? Compare with the theoretical standard deviation (see below). *Report* the result.
4. Does the distribution of μ_i look normal? *Plot* a histogram and compare with the theoretical pdf (see below).

3.1 ...basic stats (the theoretical results)

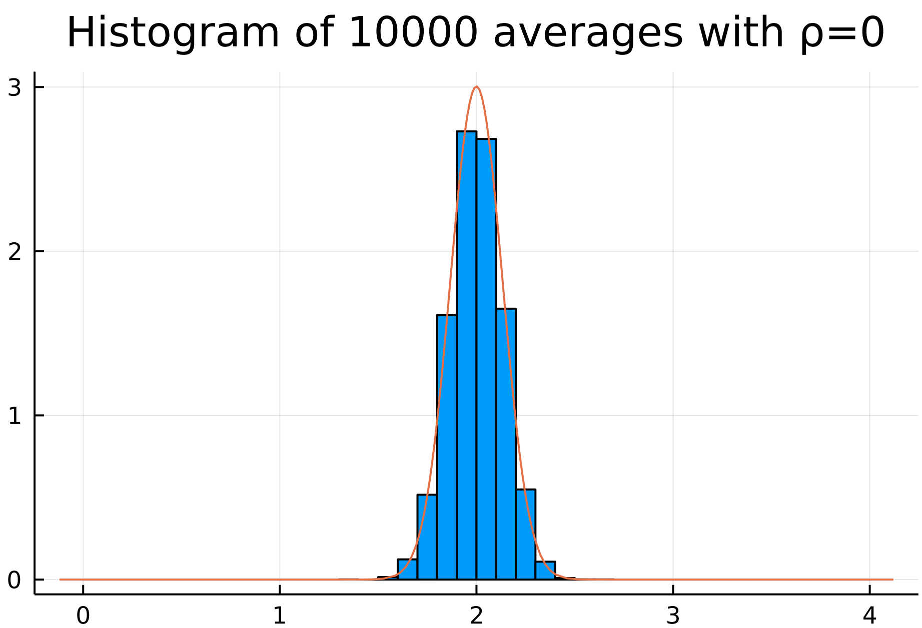
says that the sample average of an iid ("independently and identically distributed") data series is normally distributed with a mean equal to the true (population) mean μ and a standard deviation equal to $s = \sigma_y/\sqrt{T}$ where σ_y is the standard deviation of y.

To compare with our simulation results, you could estimate σ_y from a single simulation with very many observations (say 10'000).

[105]:
```
M = 10_000

(T, 1, , ) = (500,0,3,2)

 i1 = fill(NaN, M)
 i1 = fill(NaN, M)

# Monte Carlo simulation
for i = 1:M
    y = SimAR1(T,  1,  ,  )
     i1[i] = mean(y)
     i1[i] = std(y)
end

# Theoretical results
y1 = SimAR1(10_000, 1, , )
 y1 = std(y1)
s1 =  y1/sqrt(T)

printmat("Average across the simulations:", mean( i1))
println("Std across the samples (with  =0) and in theory:")
printmat(["simulations", std( i1)], ["theory", s1])
```

```
Average across the simulations:     2.001

Std across the samples (with  =0) and in theory:
simulations    theory
     0.134     0.133
```

[106]:
```
histogram( i1,bins=0:0.1:4,normalize=true,legend=false,title="Histogram of␣
 ↪10000 averages with  =0")
plot!( i1->pdf(Normal( , s1),  i1))
```



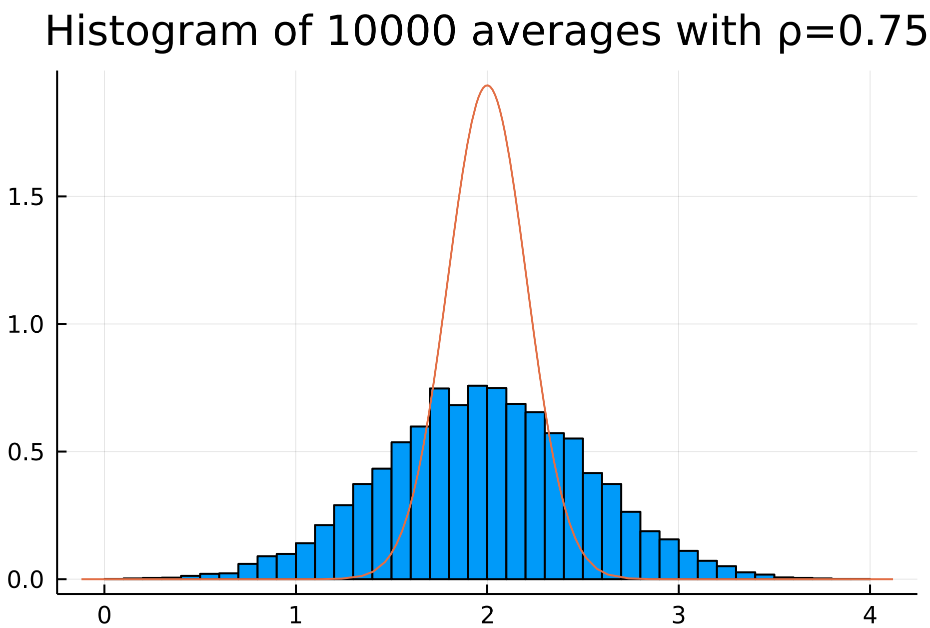
4 Task 3

Redo task 2, but now use `=0.75` (the other parameters are unchanged).

[107]:
```
M = 10_000

(T, 2, , ) = (500,0.75,3,2)

 i2 = fill(NaN, M)
 i2 = fill(NaN, M)

# Monte Carlo simulation
for i = 1:M
    y = SimAR1(T,  2,  ,  )
     i2[i] = mean(y)
     i2[i] = std(y)
end

# Theoretical results
y2 = SimAR1(10_000, 2, , )
 y2 = std(y2)
s2 =  y2/sqrt(T)

printmat("Average across the simulations:", mean( i2))
println("Std across the samples (with  =0.75) and in theory:")
printmat(["simulations", std( i2)], ["theory", s2])
```

```
Average across the simulations:     1.989

Std across the samples (with  =0.75) and in theory:
simulations    theory
     0.534     0.206
```

[108]:
```
histogram( i2,bins=0:0.1:4,normalize=true,legend=false,title="Histogram of␣
↪10000 averages with  =0.75")
plot!( i2->pdf(Normal( , s2),  i2))
```



5 Task 4

You decide to test the hypothesis that $\mu = 2$. Your decision rule is

- reject the hypothesis if $|(\mu_i - 2)/s| > 1.645$ with $s = \sigma_y/\sqrt{T}$

With this decision rule, you are clearly assuming that the theoretical result (definition of s) is correct.

Estimate both μ_i and σ_y from each sample.

In what fraction of the M simulation do you reject your hypothesis when $\rho = 0$ and when $\rho = 0.75$? For the other parameters, use `(T, , ) = (500,3,2)` (same as before).

[109]:
```
(T, , ) = (500,3,2)

rejection1 = 0
```

```
rejection2 = 0

# Count how many times we reject the hypothesis
for i = 1:M
    # rejections for   = 0
    s1 =  i1[i]/sqrt(T)
    if abs(( i1[i] - 2)/s1) > 1.645
        rejection1 += 1
    end

    # rejections for   = 0.75
    s2 =  i2[i]/sqrt(T)
    if abs(( i2[i] - 2)/s2) > 1.645
        rejection2 += 1
    end
end

println("Frequency of rejections:")
printmat(["with  =0 ", rejection1 / (M)], ["with  =0.75 ", rejection2 / (M)])
```

```
Frequency of rejections:
 with  =0 with  =0.75
     0.098     0.536
```