Project I: Random Processes

Random variables, stationarity & ergodicity

Simon Andreas Bjørn

February 7, 2023

» 1. Random variables

Importing some statistic and plotting modules

```julia
using Printf
using Statistics
using NaNStatistics
using Plots
```

» 1.1 Random variables - uniform PDF

We start by sampling the uniform destribution on $[0, 1)$, with $N = 10000$ samples.

```
N = 10_000
x = rand(N)
```

» 1.1 Random variables - uniform PDF

We start by sampling the uniform destribution on $[0, 1)$, with $N = 10000$ samples.

```
N = 10_000
x = rand(N)
```

We then find the first order momentum $m_x = E\{x\}$ and the second order central moment $\sigma_x^2 = E\{(x - m_x)^2\}$

```
# Evaluate E{x} and E{(x-mₓ)²} by definition
mₓ  = sum(x*1/N)
σₓ² = sum((x.-mₓ).^2 * 1/N)

# Evaluate E{x} and E{(x-mₓ)²} by builtin functions
mₓ_builtin = mean(x)
σₓ²_builtin = std(x)^2
```

» Calculation differences

Calculating the difference between the builtin functions and our direct evaluation, we see that the the two methods are similar.

```
@printf "mₓ_err: %.4f" abs(mₓ-mₓ_builtin)
# -> mₓ_err: 0.0000
@printf "σₓ²_err: %.4f" abs(σₓ²-σₓ²_builtin)
# -> σₓ²_err: 0.0000
```

» Calculating the PDF

We then calculate the pdf using `histcounts`, and normalize the PDF.

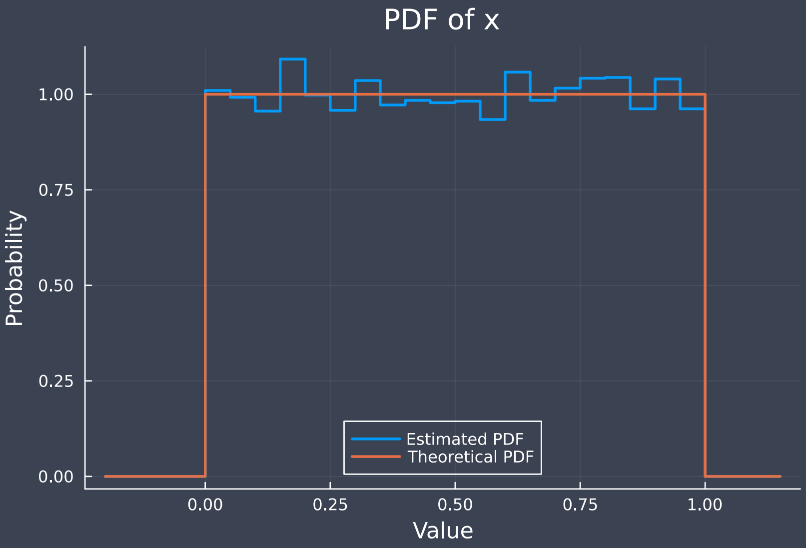
```
step_size = 0.05
edges = -0.2:step_size:1.2
bins = histcounts(x, edges)

# Normalize bins to get PDF
pdf = bins/sum(bins*step_size)

# Theoretical PDF
theoretical_pdf = (x) -> Float64(0 <= x < 1)
```

» Code for plotting the PDF

And we plot the result

```
plot(
    edges[1:end-1],
    pdf,
    seriestype=:steppost,
    label="Estimated PDF"
)
    plot!(
    edges[1:end-1],
    theoretical_pdf.(edges[1:end-1]),
    seriestype=:steppost, label="Theoretical PDF"
)
    plot!(
    legend=:bottom,
    background_color=:transparent,
    foreground_color=:white
)
title!("PDF of x");xlabel!("Value");
ylabel!("Probability");savefig("uniform_pdf.svg")
```

» Plot of the PDF



» Ensuring properties

We want to verify if our PDF fulfill the 2 following properties:

1. $\forall \alpha,\ f_x(\alpha) \geq 0$
   ```
   reduce(&, pdf .>= 0) # -> true
   ```
2. $\int_{-\infty}^{\infty} f_x(\alpha)\, \mathrm{d}\alpha = 1$
   ```
   abs(sum(pdf*step_size)-1.0) < 1e-12 # -> true
   ```

» Deriving the theoretical mean [m_x]

» Deriving the theoretical mean [m_x]

We start at the definition of the mean value which is given by

$$m_x = E\{x\} = \int_{-\infty}^{\infty} \alpha f_x(\alpha) \mathsf{d}\alpha \tag{1}$$

» Deriving the theoretical mean [m_x]

We start at the definition of the mean value which is given by

$$m_x = E\{x\} = \int_{-\infty}^{\infty} \alpha f_x(\alpha) \mathrm{d}\alpha \tag{1}$$

We then use the theoretical probability density function of a uniform distribution which is given by

$$f_x(\alpha) = \begin{cases} 0, & \alpha < a \\ \frac{1}{b-a}, & a \leq \alpha \leq b \\ 0, & \alpha > b \end{cases}$$

» Deriving the theoretical mean [m_x]

With $a = 0, b = 1$, we get the theoretical PDF

$$f_x(\alpha) = \begin{cases} 0, & \alpha < 0 \\ 1, & 0 \leq \alpha \leq 1 \\ 0, & \alpha > 1 \end{cases}$$

» Deriving the theoretical mean [m_x]

With $a = 0, b = 1$, we get the theoretical PDF

$$f_x(\alpha) = \begin{cases} 0, & \alpha < 0 \\ 1, & 0 \leq \alpha \leq 1 \\ 0, & \alpha > 1 \end{cases}$$

We simplify eq.1, and get the integral

$$m_x = E\{x\} = \int_0^1 \alpha \ \mathsf{d}\alpha$$

» Deriving the theoretical mean [m_x]

With $a = 0, b = 1$, we get the theoretical PDF

$$f_x(\alpha) = \begin{cases} 0, & \alpha < 0 \\ 1, & 0 \leq \alpha \leq 1 \\ 0, & \alpha > 1 \end{cases}$$

We simplify eq.1, and get the integral

$$m_x = E\{x\} = \int_0^1 \alpha \ \mathsf{d}\alpha$$

We simply solve this and get the mean value to be

$$m_x = \frac{1}{2}\left[\alpha^2\right]_0^1 = \frac{1}{2}$$

» Deriving the theoretical variance [σ_x^2]

» Deriving the theoretical variance [σ_x^2]

Again, we start at the definition

$$\sigma_x^2 = E\left\{(x - m_x)^2\right\} = \int_{-\infty}^{\infty} (\alpha - m_x) f_x(\alpha) \mathsf{d}\alpha \qquad (2)$$

» Deriving the theoretical variance [σ_x^2]

Again, we start at the definition

$$\sigma_x^2 = E\left\{(x - m_x)^2\right\} = \int_{-\infty}^{\infty} (\alpha - m_x) f_x(\alpha) \mathsf{d}\alpha \qquad (2)$$

We do the same simplification as with the mean value, and and fill in for m_x. This gives the variance

$$\begin{aligned} \sigma_x^2 &= \int_0^1 \left(\alpha - \frac{1}{2}\right)^2 \ \mathsf{d}\alpha \\ &= \int_0^1 \alpha^2 - \alpha + \frac{1}{4} \ \mathsf{d}\alpha \\ &= \left[\frac{1}{3}\alpha^3 - \frac{1}{2}\alpha^2 + \frac{1}{4}\alpha\right]_0^1 \\ &= \frac{1}{12} \approx 0.08334 \end{aligned}$$

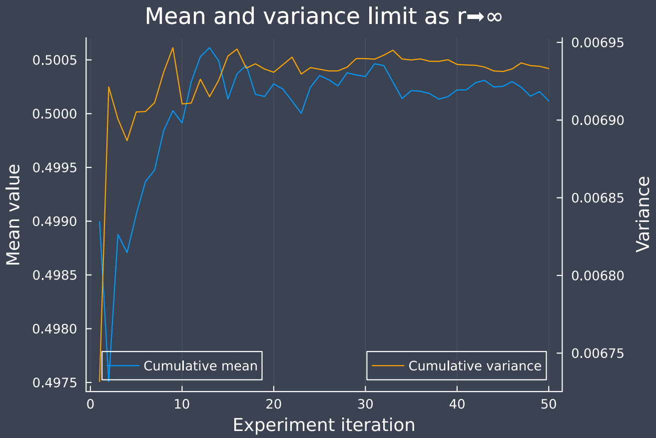
» Repeating the uniform experiment

We repeat the experiment, find the mean and variance of each experiment and accumulate the results.

```
r = 50 # Repeat r times
xs = rand(Float64, (N,r))
ms = mean(xs, dims=1)
σs = std(xs, dims=1).^2
# Accumulate the results
i_mean = [mean(ms[1:i]) for i in 1:length(ms)]
i_std = [mean(σs[1:i])^2 for i in 1:length(s)]
# Plot the results
plot(1:r, i_mean, label="Cumulative mean",
    legend=:bottomleft, xlabel="Experiment iteration",
    ylabel="Mean value")
plot!( twinx(), i_std, label="Variance",
    legend=:bottomright, color=:orange, ylabel="Variance")
plot!(foreground_color=:white,
    background_color=:transparent)
savefig!("repeated_experiment_uniform.svg")
```

» Repeating the uniform experiment

We can see that the mean and variance approach some values as we perform more experiments



» 1.2 Gaussian PDF

We perform the same experiment, but sample using `randn`.

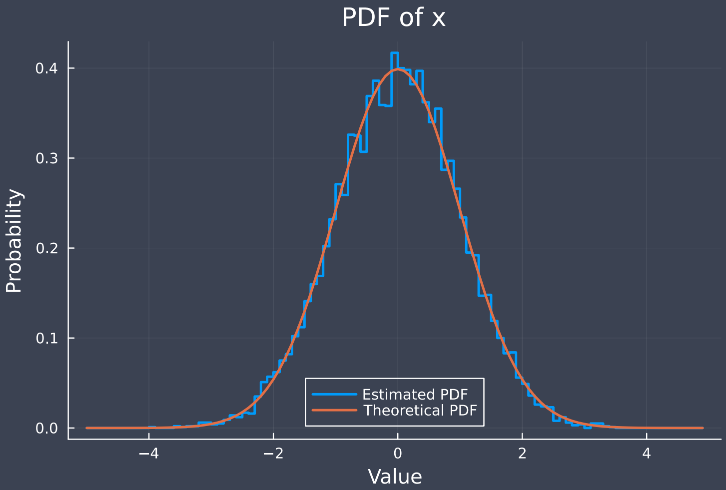
```julia
N = 10_000
x = randn(N)
mₓ  = sum(x*1/N)
σₓ² = sum((x.-mₓ).^2 * 1/N)

step_size = 0.1
edges = -5.0:step_size:5.0
bins = histcounts(x, edges)
# Normalize PDF
pdf = bins/sum(bins*step_size)
theoretical_pdf = (x) -> 1.0/√(2π)*exp(-x.^2 ./ 2)
```

» 1.2 Plotting code

We then plot the results in a similar manner as 1.1.

```julia
plot( edges[1:end-1], pdf, seriestype=:steppost,
     label="Estimated PDF", linewidth=2)

plot!(edges[1:end-1], theoretical_pdf.(edges[1:end-1]),
     label="Theoretical PDF", linewidth=2)

plot!(legend=:bottom, background_color=:transparent,
     foreground_color=:white)

title!("PDF of x");xlabel!("Value");
ylabel!("Probability");
```

» 1.2 Gaussian PDF plot



» 1.3 Central Limit Theorem

Again, we sample `rand` as according to the task. We have have the following code.

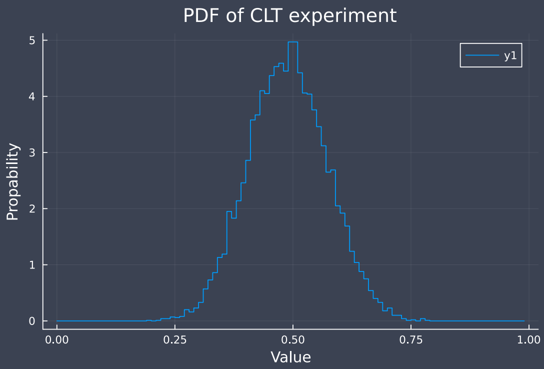
```julia
N = 10_000
x = rand(Float64, (12, N))
ms = mean(x, dims=1)

step_size = 0.01
edges = 0.0:step_size:1.0
bins = histcounts(ms, edges)

pdf = bins/sum(bins*step_size)
sum(pdf*step_size)

mₓ = mean(ms)
σₓ = std(ms)^2

@printf "mₓ: %.3f" mₓ  # -> mₓ: 0.499
@printf "σₓ²: %.3f" σₓ # -> σₓ²: 0.007
```

We plot the estimated PDF with the following code

```
plot(edges[1:end-1], pdf, seriestype=:steppre)
plot!(background_color=:transparent,foreground_color=:white)
xlabel!("Value"); ylabel!("Propability")
title!("PDF of CLT experiment")
savefig("clt_pdf.svg")
```

We plot the estimated PDF with the following code

```
plot(edges[1:end-1], pdf, seriestype=:steppre)
plot!(background_color=:transparent,foreground_color=:white)
xlabel!("Value"); ylabel!("Propability")
title!("PDF of CLT experiment")
savefig("clt_pdf.svg")
```



» Theoretical mean of CTF experiment [m_z]

We start by noting the random variable for our experiment can be defined as $Z = \frac{1}{12}\sum_{i=1}^{12} X_i$.

» Theoretical mean of CTF experiment [m_z]

We start by noting the random variable for our experiment can be defined as $Z = \frac{1}{12}\sum_{i=1}^{12} X_i$.

Then

$$E\left\{Z\right\} = E\left\{\frac{1}{12}\sum_{i=1}^{12} X_i\right\} = \frac{1}{12}E\left\{\sum_{i=1}^{12} X_i\right\} = \frac{1}{12}\sum_{i=1}^{12} E\left\{X_i\right\}$$

» Theoretical mean of CTF experiment [m_z]

We start by noting the random variable for our experiment can be defined as $Z = \frac{1}{12}\sum_{i=1}^{12} X_i$.

Then

$$E\{Z\} = E\left\{\frac{1}{12}\sum_{i=1}^{12} X_i\right\} = \frac{1}{12}E\left\{\sum_{i=1}^{12} X_i\right\} = \frac{1}{12}\sum_{i=1}^{12} E\{X_i\}$$

As we know that $E\{X_i\} = \frac{1}{2}$, we have

$$m_z = E\{Z\} = \frac{1}{12}\sum_{i=1}^{12}\frac{1}{2} = \frac{1}{2}$$

» Theoretical variance of CTF experiment [σ_z^2]

We want to find the variance $\sigma_x^2 = E\left\{(Z - \bar{Z})^2\right\}$, and using the definition from the previous slide, we have

$$E\left\{(Z - E\left\{Z\right\})^2\right\} = E\left\{\left(\frac{1}{12}\sum_{i=1}^{12} X_i - E\left\{\frac{1}{12}\sum_{i=1}^{12} X_i\right\}\right)^2\right\}$$

» Theoretical variance of CTF experiment [σ_z^2]

We want to find the variance $\sigma_x^2 = E\left\{(Z-\bar{Z})^2\right\}$, and using the definition from the previous slide, we have

$$E\left\{(Z-E\{Z\})^2\right\} = E\left\{\left(\frac{1}{12}\sum_{i=1}^{12} X_i - E\left\{\frac{1}{12}\sum_{i=1}^{12} X_i\right\}\right)^2\right\}$$

We simplify, using

$$E\{X+Y\} = E\{X\} + E\{Y\} \tag{3}$$

and get the following

$$E\left\{(Z-E\{Z\})^2\right\} = E\left\{\left(\frac{1}{12}\sum_{i=1}^{12} X_i - \frac{1}{12}\sum_{i=1}^{12} E\{X_i\}\right)^2\right\}$$

» Theoretical variance of CTF experiment [σ_z^2]

We factorize and combine our sums to get

$$E\left\{(Z - E\{Z\})^2\right\} = \left(\frac{1}{12}\right)^2 E\left\{\left(\sum_{i=1}^{12} X_i - E\{X_i\}\right)^2\right\}$$

» Theoretical variance of CTF experiment [σ_z^2]

We factorize and combine our sums to get

$$E\left\{(Z - E\{Z\})^2\right\} = \left(\frac{1}{12}\right)^2 E\left\{\left(\sum_{i=1}^{12} X_i - E\{X_i\}\right)^2\right\}$$

Expanding the binomial, we get

$$\begin{aligned} &= \frac{1}{12^2}\sum_{i=1}^{12} E\left\{(X_i - E\{X_i\})^2\right\} \\ &+ 2\sum_{\substack{i=1\\ j=2\\ i<j}}^{j\to 12} E\left\{(X_i - E\{X_i\})(X_j - E\{X_j\})\right\} \end{aligned}$$

Which almost looks good, but there is an ugly term from the binomial expansion which ruins the fun.

» Theoretical variance of CTF experiment [σ_z^2]

There is luckely an easy clean-up for this. We know that our separate observations X_i are all independent. This means that each $(X_i - E\{X_i\})$ for $i = 1, 2, \ldots, 12$ are independent.

» Theoretical variance of CTF experiment [σ_z^2]

There is luckely an easy clean-up for this. We know that our separate observations X_i are all independent. This means that each $(X_i - E\{X_i\})$ for $i = 1, 2, \ldots, 12$ are independent. We can then use the multiplicative property

Multiplicative Property

For two independent variables X and Y, the expectation of their product is equal to the product of the expectation of each variable.

$$E\{XY\} = E\{X\}E\{Y\}$$

» Theoretical variance of CTF experiment [σ_z^2]

There is luckely an easy clean-up for this. We know that our separate observations X_i are all independent. This means that each $(X_i - E\{X_i\})$ for $i = 1, 2, \ldots, 12$ are independent. We can then use the multiplicative property

Multiplicative Property

For two independent variables X and Y, the expectation of their product is equal to the product of the expectation of each variable.

$$E\{XY\} = E\{X\}E\{Y\}$$

This allows us to rewrite the additional binomial term as

$$E\{X_i - E\{X_i\}\} \cdot E\{X_j - E\{X_j\}\}$$

» Theoretical variance of CTF experiment [σ_z^2]

We then utilize e.q.(3) on each factor of the binomial term, and we see that

$$E\left\{X_i - E\left\{X_i\right\}\right\} = E\left\{X_i\right\} - E\left\{E\left\{X_i\right\}\right\} = 0$$
$$\Downarrow$$
$$2\sum_{\substack{i=1\\j=2\\i<j}}^{j\to 12} E\left\{X_i - E\left\{X_i\right\}\right\} E\left\{X_j - E\left\{X_j\right\}\right\} = 0$$

And thus the big bad evil term dissapered.

» Theoretical variance of CTF experiment [σ_z^2]

Now that we know the additional term equates to 0 and we know $E\left\{\left(X_i - E\left\{X_i\right\}\right)^2\right\} = \frac{1}{12}$, we insert into our equation and get values into our equation

$$E\left\{\left(Z - E\left\{Z\right\}\right)^2\right\} = \frac{1}{12^2}\sum_{i=1}^{12} E\left\{\left(X_i - E\left\{X_i\right\}\right)^2\right\}$$

$$\sigma_z^2 = \frac{1}{12^2}\sum_{i=1}^{12}\frac{1}{12} = \frac{1}{144} \approx 0.0069$$

» Comparing the theoretical and estimated values

We got our theoretical values to be $m_z = 0.5,\ \sigma_z^2 = 0.0069$. Comparing this to our result from earlier, we see that our estimated values of $m_z = 0.499,\ \sigma_z^2 = 0.007$ is pretty much spot on.

We then plot a gaussian curve over our previous plot with these values using the following code

```
theoretical_pdf = (x) -> 1.0/√(2π*σₓ²)*exp(-(x-mₓ).^2 ./ (2σₓ²))
plot(edges[1:end-1], pdf, seriestype=:steppost, label="Estimate")
plot!(edges[1:end-1], theoretical_pdf.(edges[1:end-1]), label="Theoretical")
plot!(background_color=:transparent,foreground_color=:white)
xlabel!("Value"); ylabel!("Propability")
title!("PDF of CLT experiment")
savefig("clt_pdf_theoretical.svg")
```

» Plot of theoretical and estimated PDF