Empirical Distribution Function

```
md"# Empirical Distribution Function"
```

Statistics III – Dr. Arturo Erdely

Consider a random sample $X_1, \ldots, X_n$ from an unknown probability distribution function F_X and let $x_1, \ldots, x_n$ be the observed values. Then the observed **empirical distribution function** (edf) is given by:

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} \mathbb{1}_{\{x_i \leq x\}}$$

```
md"Consider a random sample $X_1,\ldots,X_n$ from an unknown probability distribution function $F_X$ and let $x_1,\ldots,x_n$ be the observed values. Then the observed **empirical distribution function** (edf) is given by:

$$F_n(x)=\frac{1}{n}\sum_{i\,=\,1}^n\mathbb{1}_{\{x_i\,\leq\,x\}}$$"
```

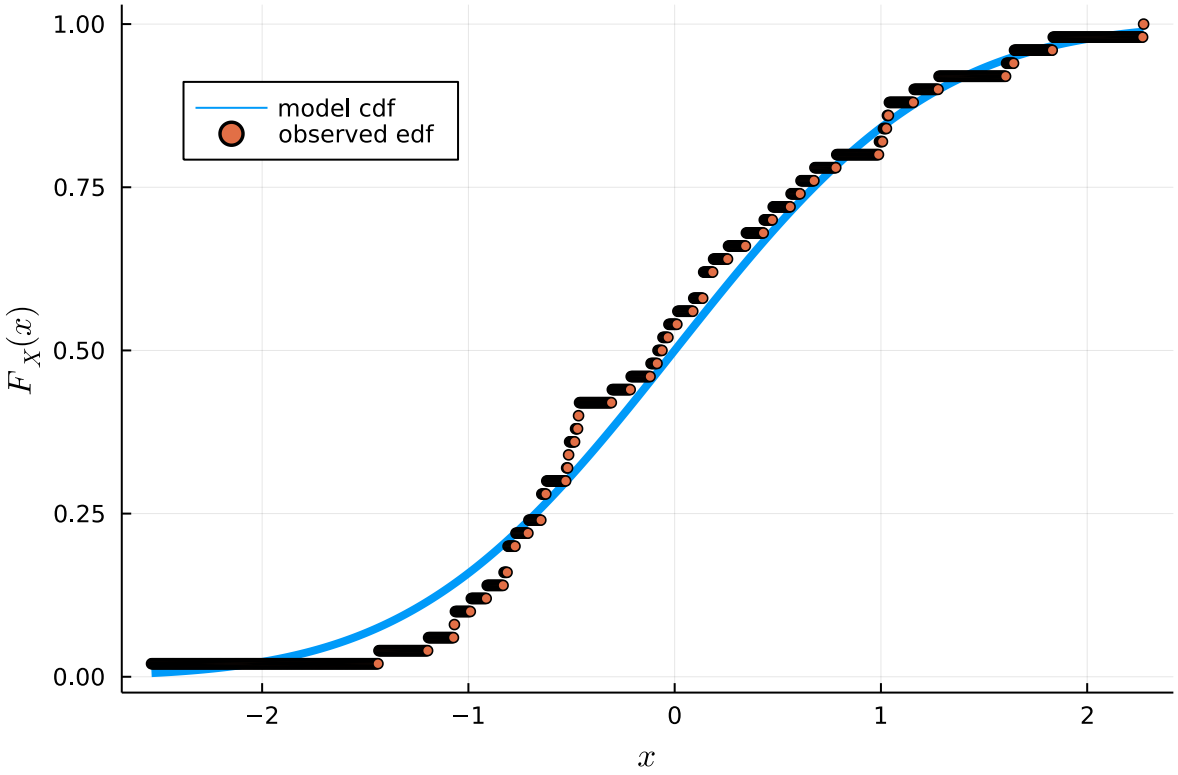
```
Fn (generic function with 1 method)
```

```
function Fn(x::Vector{<:Real}, xobs::Vector{<:Real})
    m = length(x)
    n = length(xobs)
    v = zeros(m)
    for k ∈ 1:m
        v[k] = count(xobs .≤ x[k]) / n
    end
    return v
end
```

```
using Distributions ✓, Plots ✓, LaTeXStrings ✓, PlutoUI ✓
```

Sample size n 50

```
md"""
Sample size $\,n$ $(@bind n Slider(10:10:1000; show_value=true))
"""
```




```
begin
    modelo = Normal(0,1)
    xobs = rand(modelo, n)
    x = collect(range(minimum(xobs), maximum(xobs), length = 1_000))
    y = Fn(x, xobs)
    plot(x, cdf(modelo, x), lw = 4, label = "model cdf", legend = (.15, .85))
    scatter!(x, y, markersize = 3, label = "observed edf")
    xaxis!(L"x")
    yaxis!(L"F_X(x)")
end
```