Trabajo final

Alumno: Hugo Marlon Fernandez Quiroz

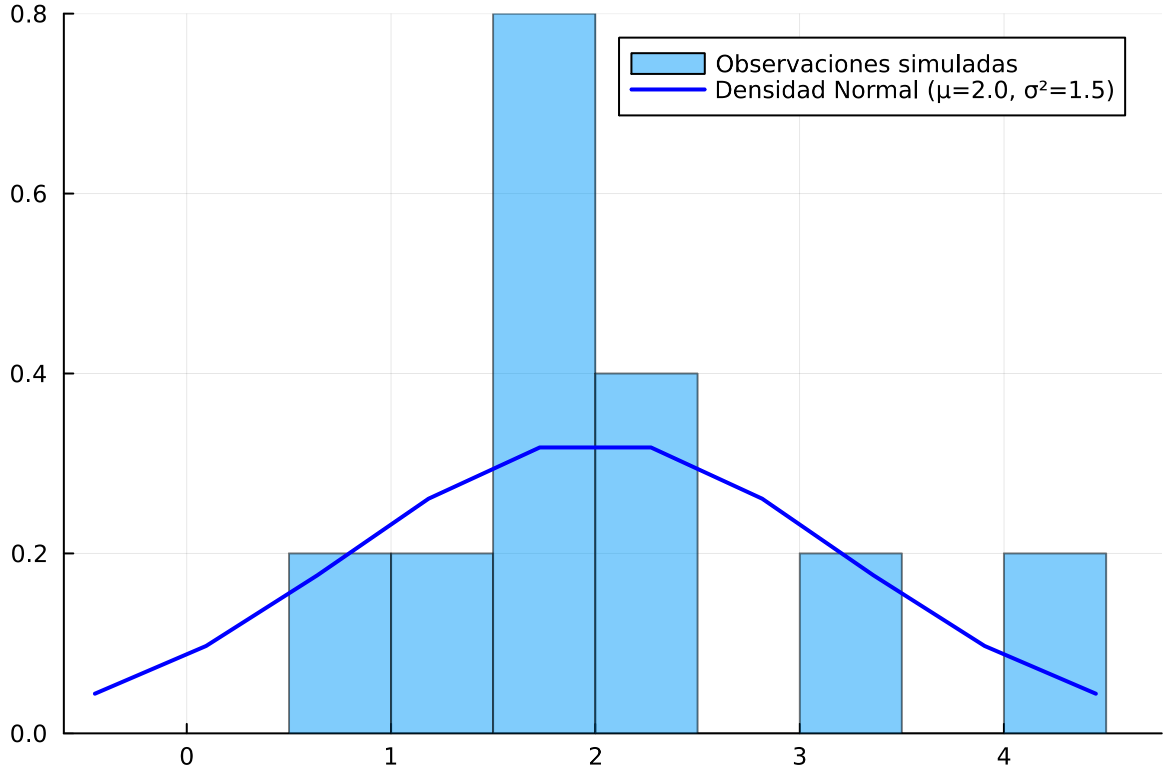
1. Considere las variables aleatorias $Y_i \sim \mathrm{Normal}(\mu, \sigma^2)$ para $i = 1, \ldots, n$, en donde el vector de parámetros es $\theta = (\mu, \sigma^2)$:

(a) Simule las observaciones para $n = 10$ para obtener $\mathbf{y} = [y_1, y_2, \ldots, y_{10}]^\top$.

```
[2.21836, 1.90046, 1.01509, 1.87073, 1.94512, 1.92636, 0.518093, 2.01743, 3.16654, 4.2580
```

```julia
begin
    # Definir parámetros
    μ = 2.0            # Media
    σ² = 1.5           # Varianza
    n = 10             # Número de observaciones

    # Simular las observaciones
    y = μ .+ sqrt(σ²) .* randn(n)
end
```



(b) Obtener la función de verosimilitud $L(\theta)$ y log-verosimilitud $\ell(\theta)$ asociada a las observaciones $y_1, y_2, \ldots, y_{10}$. Interprete la forma que tenga la función de verosimilitud.

Dado que las observaciones $y_1, y_2, \ldots, y_{10}$ siguen una distribución normal Y_i ~ Normal(μ, σ^2), la *función de verosimilitud* $L(\theta)$ se define como:

$$L(\mu, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_i - \mu)^2}{2\sigma^2}\right)$$

Donde θ = (μ, σ^2) es el vector de parámetros.

La log-verosimilitud $\ell(\theta)$ es el logaritmo natural de la función de verosimilitud:

$$\ell(\mu, \sigma^2) = \log L(\mu, \sigma^2) = -\frac{n}{2}\log(2\pi\sigma^2) - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - \mu)^2$$

```
begin
    function verosimilitud(y, μ, σ²)
        n = length(y)
        (1 / ((2 * π * σ²)^(n/2))) * exp(-sum((y .- μ).^2) / (2 * σ²))
    end

    # Log-verosimilitud
    function log_verosimilitud(y, μ, σ²)
        n = length(y)
        -(n/2) * log(2 * π * σ²) - sum((y .- μ).^2) / (2 * σ²)
    end

    println("Verosimilitud: ", verosimilitud(y, μ, σ²))
    println("Log-verosimilitud: ", log_verosimilitud(y, μ, σ²))
end
```

```
Verosimilitud: 5.286492839449081e-7
Log-verosimilitud: -14.452940604266272
```

c) Obtener la función de score S(θ) y la información de Fisher $I_E(\theta)$

La *función score* es la derivada de la log-verosimilitud con respecto a los parámetros μ y σ² :

Para μ:

$$S_\mu(\mu, \sigma^2) = \frac{\partial \ell(\mu, \sigma^2 \mid y)}{\partial \mu} = \frac{1}{\sigma^2} \sum_{i=1}^{n} (y_i - \mu)$$

Para σ² :

$$S_{\sigma^2}(\mu, \sigma^2) = \frac{\partial \ell(\mu, \sigma^2 \mid y)}{\partial \sigma^2} = -\frac{n}{2\sigma^2} + \frac{1}{2\sigma^4} \sum_{i=1}^{n} (y_i - \mu)^2$$

La *información de Fisher* es la matriz de varianza-covarianza inversa de los estimadores de máxima verosimilitud, cuyas entradas son los valores esperados de los productos de las derivadas de la log-verosimilitud:

Para μ, σ²:

$$\mathcal{I}(\mu, \sigma^2) = \begin{pmatrix} \frac{n}{\sigma^2} & 0 \\ 0 & \frac{n}{2\sigma^4} \end{pmatrix}$$

```
begin
    function score(y, μ, σ²)
        n = length(y)
        dL_dμ = sum(y .- μ) / σ²
        dL_dσ2 = -(n / (2 * σ²)) + sum((y .- μ).^2) / (2 * σ²^2)
        return dL_dμ, dL_dσ2
    end

    # Información de Fisher
    function fisher_information(y, μ, σ²)
        n = length(y)
        I_μμ = n / σ²
        I_μσ2 = 0
        I_σ2σ2 = n / (2 * σ²^2)
        return I_μμ, I_μσ2, I_σ2σ2
    end

    println("Score: ", score(y, μ, σ²))
    println("Información de Fisher: ", fisher_information(y, μ, σ²))

end
```

```
Score: (0.5574855579592516, -1.1758468455475186)
Información de Fisher: (6.666666666666667, 0, 2.2222222222222223)
```

(d) Dado un conjunto de observaciones simuladas, obtenga intervalos de confianza utilizando la distribución asintótica del estimador de máxima verosimilitud.

Los estimadores de máxima verosimilitud para μ y σ² son:

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} y_i, \quad \hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{\mu})^2$$

Para el intervalo de confianza de μ:

$$\hat{\mu} \pm z_{\alpha/2} \cdot \sqrt{\frac{\hat{\sigma}^2}{n}}$$

donde $z_{\alpha/2}$ es el valor crítico de la distribución normal estándar.

Para σ², el intervalo de confianza es:

$$\left(\frac{(n-1)\hat{\sigma}^2}{\chi^2_{\alpha/2,n-1}}, \frac{(n-1)\hat{\sigma}^2}{\chi^2_{1-\alpha/2,n-1}}\right)$$

donde $\chi^2_{\alpha,k}$ es el valor crítico de la distribución chi-cuadrado con k grados de libertad.

```
begin
    μ_hat = mean(y)
    σ2_hat = var(y)

    # Intervalos de confianza
    α = 0.05  # Nivel de significancia
    z = quantile(Normal(0, 1), 1 - α/2)

    conf_μ = (μ_hat - z * sqrt(σ2_hat / n), μ_hat + z * sqrt(σ2_hat / n))
    conf_σ2 = (σ2_hat * (n - 1) / quantile(Chisq(n - 1), 1 - α/2), σ2_hat * (n - 1)
    / quantile(Chisq(n - 1), α/2))

    println("Intervalo de confianza para μ: ", conf_μ)
    println("Intervalo de confianza para σ2: ", conf_σ2)


end
```

```
Intervalo de confianza para μ: (1.4422102862101966, 2.725035381177578)
Intervalo de confianza para σ2: (0.5066960556904075, 3.569396715531712)
```

(e) Dado un conjunto de observaciones simuladas, obtenga intervalos de confianza utilizando la devianza.

La devianza es una medida de la calidad del ajuste del modelo. Se define como:

$$D = -2 \cdot \ell(\hat{\mu}, \hat{\sigma}^2 \mid y)$$

Para construir intervalos de confianza usando la devianza, se compara con la distribución χ^2 asintótica bajo ciertas condiciones del modelo.

```
begin
    function devianza(y, μ_hat, σ2_hat)
        -2 * log_verosimilitud(y, μ_hat, σ2_hat)
    end

    # Intervalos de confianza con la devianza (asume chi-cuadrado)
    conf_devianza = [devianza(y, μ_hat, σ2_hat) + quantile(Chisq(1), α/2),
    devianza(y, μ_hat, σ2_hat) + quantile(Chisq(1), 1 - α/2)]
    println("Intervalo de confianza usando la devianza: ", conf_devianza)

end
```

```
Intervalo de confianza usando la devianza: [28.065433121083213, 33.08833723
9280925]
```

(f) Simulando otra muestra $x_1, x_2, \ldots, x$ n con distinta media $(Normal(\mu_2, \sigma^2))$. Realizar la prueba de hipótesis para $\mu_1 = \mu_2$ utilizando la prueba de razones de verosimilitud.

Para probar la hipótesis $H_0 : \mu_1 = \mu_2$ contra $H_1 : \mu_1 \neq \mu_2$, usamos la *razón de verosimilitudes*:

$$\Lambda = -2\left[\ell(\hat{\mu}, \hat{\sigma}^2 \mid y) - (\ell(\hat{\mu}_1, \hat{\sigma}_1^2 \mid y_1) + \ell(\hat{\mu}_2, \hat{\sigma}_2^2 \mid y_2))\right]$$

```
begin
    # segunda muestra
    x_new = rand(Normal(0, sqrt(σ²)), n)

    # Hipótesis nula: μ1 = μ2
    logL0 = log_verosimilitud([y; x_new], mean([y; x_new]), σ2_hat)
    # Hipótesis alternativa: μ1 ≠ μ2
    logL1 = log_verosimilitud(y, mean(y), σ2_hat) + log_verosimilitud(x_new,
    mean(x_new), σ2_hat)

    # Estadístico de prueba
    LR = -2 * (logL0 - logL1)

    # P-valor
    p_value = 1 - cdf(Chisq(1), LR)
    println("Razón de verosimilitudes: ", LR)
    println("P-valor: ", p_value)

    significancia = 0.05
    if p_value < significancia
        println("Se rechaza la hipótesis nula al 95% de significancia")
    else
        println("Se acepta la hipótesis nula")
    end
end
```

```
Razón de verosimilitudes: 13.00915122505365
P-valor: 0.0003099724098317713
Se rechaza la hipótesis nula al 95% de significancia
```

2. Continuando con el problema anterior, asuma los priors μ ~ $Normal(\mu_0, \sigma^2{}_0)$ y σ² ~ InverseGamma(a, b).

(a) Dada la muestra simulada y, obtener la distribución aposteriori condicional π(μ | σ² , y).

Dada la muestra $\mathbf{y} = (y_1, y_2, \ldots, y_n)$, y suponiendo que σ² es conocida:

1. *Modelo*: Supongamos que los datos son independientes y normalmente distribuidos: $y_i \mid \mu, \sigma^2 \sim \text{Normal}(\mu, \sigma^2), \quad i = 1, 2, \ldots, n$

Es decir, la función de verosimilitud es: $p(\mathbf{y} \mid \mu, \sigma^2) = \prod_{i=1}^n \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(y_i-\mu)^2}{2\sigma^2}\right)$

Esto se puede escribir como: $p(\mathbf{y} \mid \mu, \sigma^2) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^n (y_i - \mu)^2\right)$

2. *Prior de μ*: El prior de μ es una distribución normal: $\mu \sim \text{Normal}(\mu_0, \sigma_0^2)$

Entonces, la función de densidad del prior es: $p(\mu) = \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left(-\frac{(\mu-\mu_0)^2}{2\sigma_0^2}\right)$

3. *Posterior de μ*: La posterior de μ dado σ^2 y los datos es proporcional al producto de la verosimilitud y el prior:

$$p(\mu \mid \sigma^2, \mathbf{y}) \propto p(\mathbf{y} \mid \mu, \sigma^2) p(\mu)$$

Sustituyendo las expresiones de la verosimilitud y el prior:

$$p(\mu \mid \sigma^2, \mathbf{y}) \propto \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-\mu)^2 - \frac{1}{2\sigma_0^2}(\mu-\mu_0)^2\right)$$

4. *Completando el cuadrado*: Simplificamos la expresión completando el cuadrado en μ:

$$p(\mu \mid \sigma^2, \mathbf{y}) \propto \exp\left(-\frac{1}{2}\left(\frac{1}{\sigma_0^2}+\frac{n}{\sigma^2}\right)(\mu-\mu_n)^2\right)$$

donde: $\mu_n = \frac{\frac{\mu_0}{\sigma_0^2}+\frac{n\bar{y}}{\sigma^2}}{\frac{1}{\sigma_0^2}+\frac{n}{\sigma^2}}, \quad \sigma_n^2 = \left(\frac{1}{\sigma_0^2}+\frac{n}{\sigma^2}\right)^{-1}$

Así, se obtiene que: $\mu \mid \sigma^2, \mathbf{y} \sim \text{Normal}(\mu_n, \sigma_n^2)$

```julia
begin
    # Parámetros a priori
    μ0 = 0.0
    σ0² = 1.0

    # Datos simulados
    y_bayes = rand(1:100, 5)
    n_bayes = length(y)
    ȳ = mean(y)

    # Dado un valor de σ²
    σ²_bayes = 1.0  # Supón que es un valor conocido o inicial

    # Parámetros de la distribución a posteriori de μ
    μ_n = (σ0² * ȳ + σ² * μ0) / (n_bayes * σ0² + σ²_bayes)
    σ_n² = (σ0² * σ²) / (n_bayes * σ0² + σ²_bayes)

    # Distribución a posteriori condicional de μ dado σ² y y
    posterior_μ = Normal(μ_n, sqrt(σ_n²))

    # Print para la sección (a)
    println("Distribución a posteriori condicional de μ | σ², y:")
    println("μ ~ Normal($μ_n, $(sqrt(σ_n²)))")

end
```

```
Distribución a posteriori condicional de μ | σ², y:
μ ~ Normal(0.1894202576085352, 0.3692744729379982)
```

(b) Dada la muestra simulada y, obtener la distribución aposteriori condicional π(σ² | μ, y).

Dada la muestra $\mathbf{y} = (y_1, y_2, \ldots, y_n)$, y suponiendo que μ es conocida:

1. *Modelo*: Como antes, los datos siguen una distribución normal: $y_i \mid \mu, \sigma^2 \sim \text{Normal}(\mu, \sigma^2)$

Entonces, la función de verosimilitud es:

$p(\mathbf{y} \mid \mu, \sigma^2) = \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^n (y_i - \mu)^2\right)$

2. *Prior de σ^2*: El prior de σ^2 es una distribución Inversa Gamma: $\sigma^2 \sim \text{InverseGamma}(a, b)$

Entonces, la función de densidad del prior es:

$p(\sigma^2) \propto (\sigma^2)^{-(a+1)} \exp\left(-\frac{b}{\sigma^2}\right)$

3. **Posterior de σ^2**: La posterior de σ^2 dado μ y los datos es proporcional al producto de la verosimilitud y el prior:

$$p(\sigma^2 \mid \mu, \mathbf{y}) \propto p(\mathbf{y} \mid \mu, \sigma^2) p(\sigma^2)$$

Sustituyendo las expresiones de la verosimilitud y el prior:

$p(\sigma^2 \mid \mu, \mathbf{y}) \propto (\sigma^2)^{-\frac{n}{2}-(a+1)} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^n (y_i - \mu)^2 - \frac{b}{\sigma^2}\right)$

4. *Reconociendo la forma*: Esta expresión es de la forma de una distribución Inversa Gamma:

$$p(\sigma^2 \mid \mu, \mathbf{y}) \sim \text{InverseGamma}\left(a_n, b_n\right)$$

donde: $a_n = a + \frac{n}{2}, \quad b_n = b + \frac{1}{2}\sum_{i=1}^n (y_i - \mu)^2$

```
begin
    # Parámetros a priori
    a = 1.0
    b = 0.0

    # Parámetros de la distribución a posteriori de σ²
    α_n = a + n/2
    β_n = b + 0.5 * sum((y .- μ).^2)

    # Distribución a posteriori condicional de σ² dado μ y y
    posterior_σ² = InverseGamma(α_n, β_n)

    # Print para la sección (b)
    println("Distribución a posteriori condicional de σ² | μ, y:")
    println("σ² ~ InverseGamma($α_n, $β_n)")

end
```

```
Distribución a posteriori condicional de σ² | μ, y:
σ² ~ InverseGamma(6.0, 4.854344597518083)
```

(c) Utilice el muestreo de Gibbs para obtener realizaciones de la distribución aposteriori $\pi(\mu, \sigma^2|y)$.

El muestreo de Gibbs es una técnica de muestreo secuencial que se basa en muestrear de las distribuciones condicionales completas:

1. *Inicialización*: Asigne valores iniciales $\mu^{(0)}$ y $\sigma^{2(0)}$.
2. *Iteración*:

- Muestree $\mu^{(t)} \sim \text{Normal}(\mu_n, \sigma_n^2)$ dado $\sigma^{2(t-1)}$.
- Muestree $\sigma^{2(t)} \sim \text{InverseGamma}(a_n, b_n)$ dado $\mu^{(t)}$.

Se repite estos pasos para obtener una cadena de muestras $\{(\mu^{(t)}, \sigma^{2(t)})\}_{t=1}^{T}$ que convergerá a la distribución conjunta posterior.

```
begin
    iterations = 10000
    burn_in = 2000
    μ_samples = zeros(iterations)
    σ²_samples = zeros(iterations)

    # Valores iniciales
    μ_samples[1] = mean(y)
    σ²_samples[1] = var(y)

    # Iteraciones del muestreo de Gibbs
    for i in 2:iterations
        # Muestreo de μ dado σ², y
        σ² = σ²_samples[i-1]
        μ_n = (σ0² * ȳ + σ² * μ0) / (n * σ0² + σ²)
        σ_n² = (σ0² * σ²) / (n * σ0² + σ²)
        μ_samples[i] = rand(Normal(μ_n, sqrt(σ_n²)))

        # Muestreo de σ² dado μ, y
        α_n = a + n/2
        β_n = b + 0.5 * sum((y .- μ_samples[i]).^2)
        σ²_samples[i] = 1 / rand(Gamma(α_n, 1/β_n))
    end

    # Print para la sección (c)
    println("Muestreo de Gibbs completado con $iterations iteraciones.")

    # Mostrar algunas realizaciones de las muestras
    println("\nAlgunas realizaciones de las muestras de μ y σ²:")
    for i in 1:5  # Imprime las primeras 5 realizaciones como ejemplo
        println("Iteración $i: μ = $(μ_samples[i + burn_in]), σ² = $(σ²_samples[i + burn_in])")
    end

end
```

```
Muestreo de Gibbs completado con 10000 iteraciones.

Algunas realizaciones de las muestras de μ y σ²:
Iteración 1: μ = 0.46643981723670713, σ² = 2.705781497725431
Iteración 2: μ = -0.21395447329451753, σ² = 5.7187304739861435
Iteración 3: μ = -0.42574792430690733, σ² = 6.656350321341007
Iteración 4: μ = -0.5551904831033352, σ² = 6.054644893483367
Iteración 5: μ = 1.3777046636507986, σ² = 1.018566872404387
```

(d) Usando realizaciones de la distribución aposteriori, proporcione inferencia puntual y por intervalos para ambos parámetros.

```julia
begin
    # Descartar el burn-in
    μ_samples_new = μ_samples[burn_in+1:end]
    σ²_samples_new = σ²_samples[burn_in+1:end]

    # Calcular la media posterior y los intervalos creíbles
    μ_mean = mean(μ_samples_new)
    σ²_mean = mean(σ²_samples_new)
    μ_cred_int = quantile(μ_samples_new, [0.025, 0.975])
    σ²_cred_int = quantile(σ²_samples_new, [0.025, 0.975])

    # Print para la sección (d)
    println("Resultados de la inferencia:")
    println("Media posterior de μ: $μ_mean")
    println("Intervalo creíble del 95% para μ: $μ_cred_int")
    println("Media posterior de σ²: $σ²_mean")
    println("Intervalo creíble del 95% para σ²: $σ²_cred_int")

end
```

```
Resultados de la inferencia:
Media posterior de μ: 0.13751794554254895
Intervalo creíble del 95% para μ: [-1.0113042499919969, 1.2483559409937413]
Media posterior de σ²: 5.074286530378358
Intervalo creíble del 95% para σ²: [1.2116938899967393, 14.144587725648709]
```