A) Robustez.

1. Consideremos una v.a. X procedente de la distribución contaminada

$$F \sim (1-\varepsilon)N(\mu,\sigma^2) + \varepsilon N(\mu, c^2\sigma^2),$$

es decir, con probabilidad $1-\varepsilon$, X se distribuye $N(\mu,\sigma^2)$ y con probabilidad ε se distribuye $N(\mu, c^2\sigma^2)$.

   (a) Calcular la esperanza y la varianza de X.

   (b) Sea $\bar{X}_1$ la media muestral calculada a partir de una muestra aleatoria de una distribución normal $N(\mu, c^2\sigma^2)$ y $\bar{X}_2$ la media muestral calculada a partir de una muestra aleatoria de la distribución contaminada. Calcular la eficiencia relativa entre las dos medias e interpretar el resultado obtenido.

2. Sea $\hat{\mu}(x_1,...,x_n)$ el M–estimador de posición dado por

$$\sum_{i=1}^{n} \psi((x_i - \hat{\mu})/\sigma) = 0$$

donde σ es conocido. Probar que tiene las siguientes propiedades

(i) $\hat{\mu}(x_1+a,...,x_n+a) = \hat{\mu}(x_1,...,x_n) + a$

(ii) $\hat{\mu}(-x_1,...,-x_n) = -\hat{\mu}(x_1,...,x_n)$

3. Sea $x_i = \mu + \sigma u_i$, $\sigma > 0$ con los u_i v.a.i.i.d. con distribución F simétrica alrededor del 0 y con esperanza finita. Probar que (i) y (ii) implican que $\hat{\mu}$ es un estimador insesgado para μ.

4. Sea $\hat{\sigma} =$MAD$(x_1,...,x_n)$ probar que

$$\hat{\sigma}(bx_1+a,...,bx_n+a) = |b|\hat{\sigma}(x_1,...,x_n)$$

5. Sea $\hat{\mu}(x_1,...,x_n)$ el M–estimador dado por

$$\sum_{i=1}^{n} \psi((x_i - \hat{\mu})/\hat{\sigma}) = 0$$

Mostrar que $\hat{\mu}$ ademas de satisfacer (i) y (ii) del problema 2 también satisface

(iii) $\hat{\mu}(bx_1,...,bx_n) = b\hat{\mu}(x_1,...,x_n)$.

6. Calcular k tal que $V(\psi_k^H, F\ ) = 1.05$, 1.10 y 1.25 cuando F es N(0,1).

7. Probar que el punto de ruptura de la mediana y del MAD es $\frac{[\frac{n}{2}]+1}{n}$ si n es impar y $\frac{1}{2}$ si n es par.

B) Ejercicios para hacer en la computadora

1. Generar muestras aleatorias de tamaño 20 de las siguientes distribuciones;

C_0: N(0,1)

C_1: Igual que C_0 cambiando la ultima observación por un outlier cuyo valor sea a) $x = 2.5$, b) $x = 10$ y c) $x = 100$

C_2: Cauchy

y calcular los siguientes estimadores:

a) media

b) mediana

c) media podada con $\alpha = 0.1$ y $\alpha = 0.25$

d) $M-$estimador con ψ_k^H, $k = 1,345$, usando como estimador de dispersión el MAD.

Explicar el comportamiento en cada caso de los distintos estimadores.

2. Hacer un programa que calcule el $M-$estimador para el modelo de posición cuando σ es desconocido.