Chapter 2: Location and Scale

Jana Osea

2.1 Location model

- **Simplest location model**: We assume that the outcome x_i of each observation depends on the "true value" μ of the unknown parameter and some random error process. A simple *location* model is

$$x_i = \mu + u_i \ (i = 1, \cdots, n) \tag{2.1}$$

- **Shift** and **Scale equivariance**: Let $\hat{\mu}$ be the sample mean and c be any constant, then $\hat{\mu}$ hold the following properties.

$$\hat{\mu}(x_1 + c, \cdots, x_n + c) = \hat{\mu}(x_1, \cdots, x_n) + c \text{ (shift equivariance)}$$

$$\hat{\mu}(cx_1, \cdots, cx_n) = c\hat{\mu}(x_1, \cdots, x_n) \text{ (scale equivariance)}$$

2.2 Formalizing departures from normality

- **Contaminated normal distributions (Tukey-Huber model)**: Let A be the event "the appratus fails" which has $P(A) = \epsilon$. Then F is is a *mixture* of G and H

$$\begin{aligned} F(t) &= P(X \leq t) = P(X \leq t|A')P(A') + P(X \leq t)|(A) \\ &= G(t)(1 - \epsilon) + H(t)\epsilon \end{aligned}$$

- *Example 1*: Assume that 95% of our observations are well-behaved, represented by $G = N(\mu, 1)$, but that 5% of the times the measuring system gives an erratic result, represented by a normal distribution with the same mean but a 10-fold increase in the standard deviation. We thus have the model with $\epsilon = 0.05$ and $H = N(\mu, 100)$. The model is

$$F = (1 - \epsilon)N(\mu, 1) + \epsilon N(\mu, 100)$$

2.3 M-estimators of location

- **Generalizing maximum likelihood estimators** We know that MLE's maximize the likelihood function

$$\hat{\mu} = \arg\max_{\mu} \; L(\mu) = \arg\max_{\mu} \prod_{i=1}^{n} f(x_i, \mu)$$

We generalize the MLE by maximizing other functions instead of the likelihood function. Given a function ρ, an *M-estimator of location* is a solution of

$$\hat{\mu} = \arg\min_{\mu} \; \sum_{i=1}^{n} \rho(x_i, \mu)$$

where ρ is a function with certain properties.

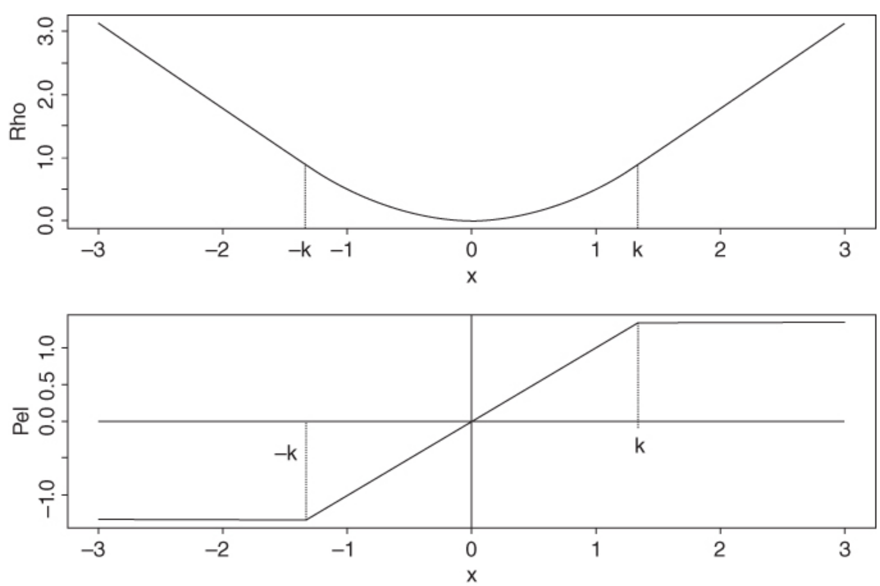
- **Huber functions**: These functions are other loss functions that we can minimize.

1. ρ-type

$$\rho_k(x) = \begin{cases} x^2 & |x| \leq k \\ 2k|x| - k^2 & |x| > k \end{cases}$$

2. ψ-type

$$\psi_k(x) = \begin{cases} x & |x| \leq k \\ sgn(x)k & |x| > k \end{cases}$$



2.4 Trimmed and Winsorized means

2.5 M-estimators of scale

2.6 Dispersion estimators

2.7 M-estimators of location with unknown dispersion

2.8 Numerical computing of M-estimators

2.9 Robust confidence intervals and tests