STT 1000 - Statistiques

Arthur Charpentier

QQ Plot

Étant donné un échantillon $\boldsymbol{y} = \{y_1, \cdots, y_n\}$, on note $\{y_{(1)}, \cdots, y_{(n)}\}$ la statistique d'ordre associée, au sens où $y_{(1)} \leq y_{(1)} \leq \cdots \leq y_{(n)}$

QQ Plot

Le QQ-plot pour un échantillon $\boldsymbol{y}$ et une loi F_0 est le nuage de points

$$\left(y_{(i)}, F_0^{-1}\left(\frac{i}{n+1}\right)\right),\ i = 1, 2, \cdots, n.$$

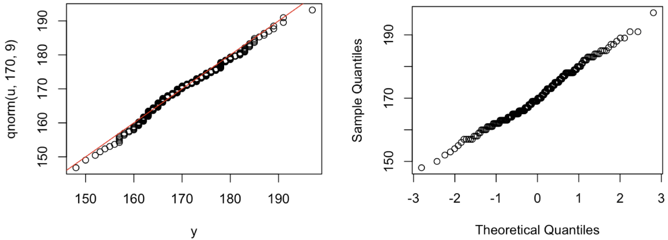
Pour rappel, si $\widehat{F}$ est la fonction de répartition empirique,

$$y_{(i)} = \widehat{F}^{-1}\left(\frac{i}{n}\right)$$

Autrement dit, on compare $\widehat{F}^{-1}$ et F_0^{-1} en des points particulier (les $i/(n+1)$ ou i/n).

QQ Plot

```
> y = sort(Davis$height)
> n = length(y)
> u = (1:n)/(n+1)
> plot(y, qnorm(u,170,9))
> qqnorm(y)
```



Note dans le cas Gaussien, si F_0 est la loi $\mathcal{N}(\mu,\sigma^2)$,

$$F_0^{-1}(u) = \mu + \sigma\Phi^{-1}(u)$$

QQ Plot

QQ Plot Gaussien

Le QQ-plot Gaussien pour un échantillon $\boldsymbol{y}$ et une loi F_0 est le nuage de points

$$\left(y_{(i)}, \Phi^{-1}\left(\frac{i}{n+1}\right)\right), \ i = 1, 2, \cdots, n.$$

Note simulation de quelques lois paramétriques, $\mathcal{N}(0,1)$, $\mathcal{S}td(3)$, $\mathcal{P}(100)$, et $LN(0,1)$, pour $n = 100$