



STK 573

Metode Grafik untuk Analisis dan Penyajian Data

Pertemuan 7

Diagnostik Plot Sebaran Peubah Kontinu Tunggal

Tim Dosen:

Prof. Dr. Khairil Anwar Notodiputro

Dr. Ir. Aji Hamim Wigena

Dr. Agus M Soleh

Introduction

- checking whether the data follow an assumed distribution
- more efficient than the EDF → the reference distribution is presented on either plot by a straight line

Introduction

- The Quantile-Quantile Plot
- The Probability Plot

The Quantile-Quantile Plot

- QQ plot ; quantile plot
- proposed by Wilk and Gnanadesikan
- the plot of two inverse distribution (or quantile) functions, $Q_1(p)$ and $Q_2(p)$, for $0 < p < 1$
- The points $\{(Q_1(p_k), Q_2(p_k))\}$ are plotted in the Cartesian coordinate plane corresponding to selected values of $\{p_k\}$ determined from an ordered random sample

- Potential values for p_k :

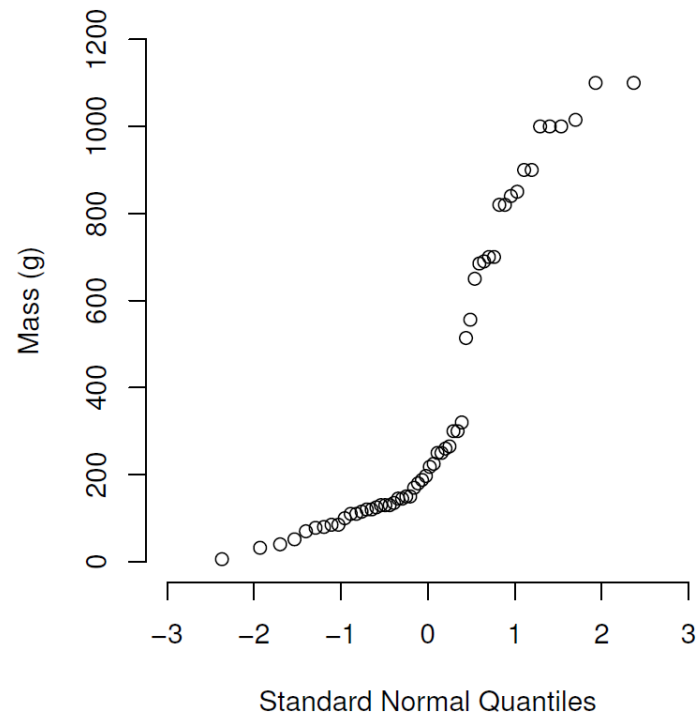
$$p_k = \frac{k - \alpha}{n - \alpha - \beta + 1}.$$

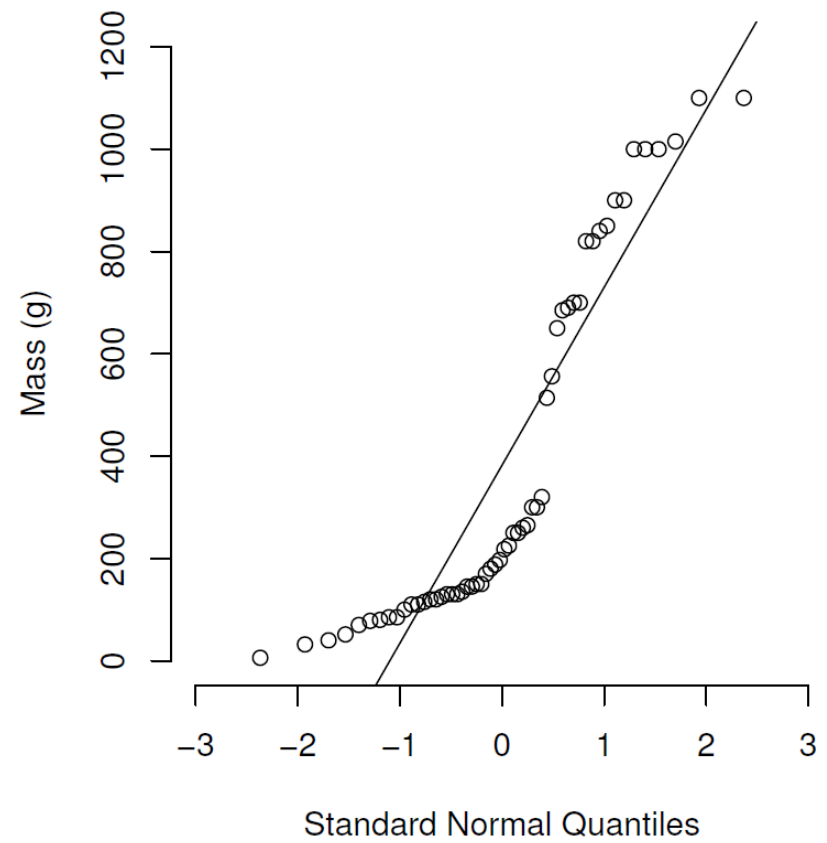
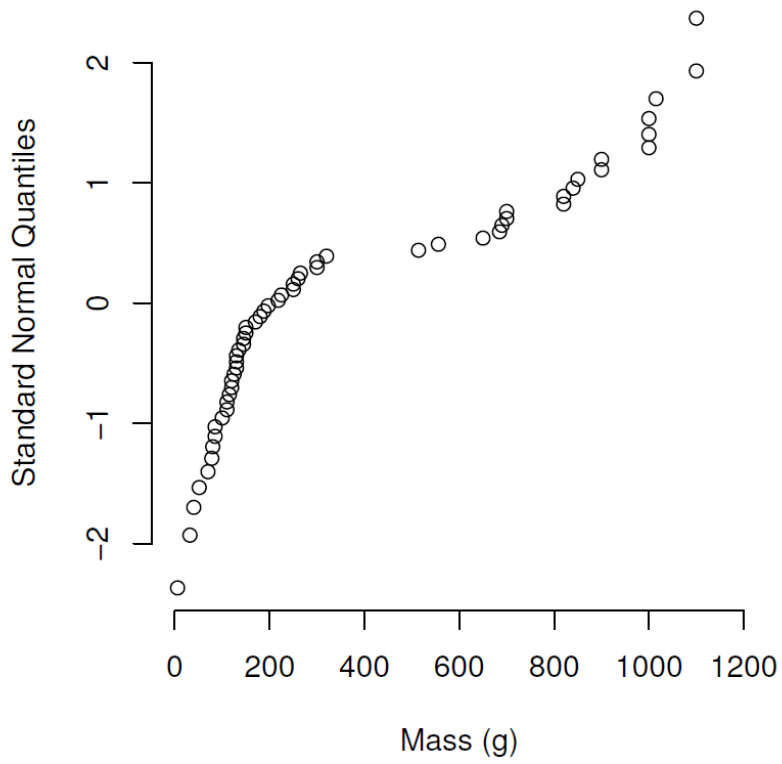
If the distribution corresponding to Q_2 is the uniform distribution function given by

$$P_2(x) = \begin{cases} 1 & \text{if } x > 1, \\ x & \text{if } 0 \leq x \leq 1, \text{ and} \\ 0 & \text{if } x < 0 \end{cases}$$

then the order statistics for the sample are plotted along the vertical axis

- The most common choices of values for Q_1 are the corresponding quantiles from the standard normal distribution $z_k = \Phi(p_k)$



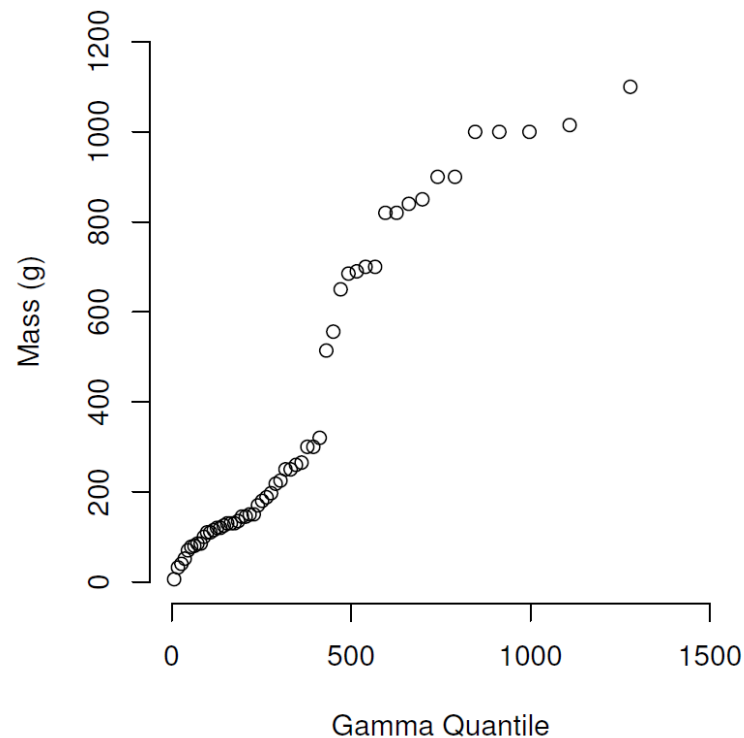


- Fungsi:
 - qqplot(); qqnorm(); qqline()
 - ppoints()
- if the data do follow a normal distribution then the points in the plot should lie nearly along a straight line
- standard procedure is to add a reference representing a normal distribution

- Normal quantile plots can be used to characterize data beyond simply checking normality:
 - If all but a few points fall on the normal reference line, then these few points may be outliers
 - If the left end of the data is above the line and the right end of the pattern is below the line, then the distribution may have short tails at both ends
 - If the left end of the data is below the line and the right end of the pattern is above the line, then the distribution may have long tails at both ends

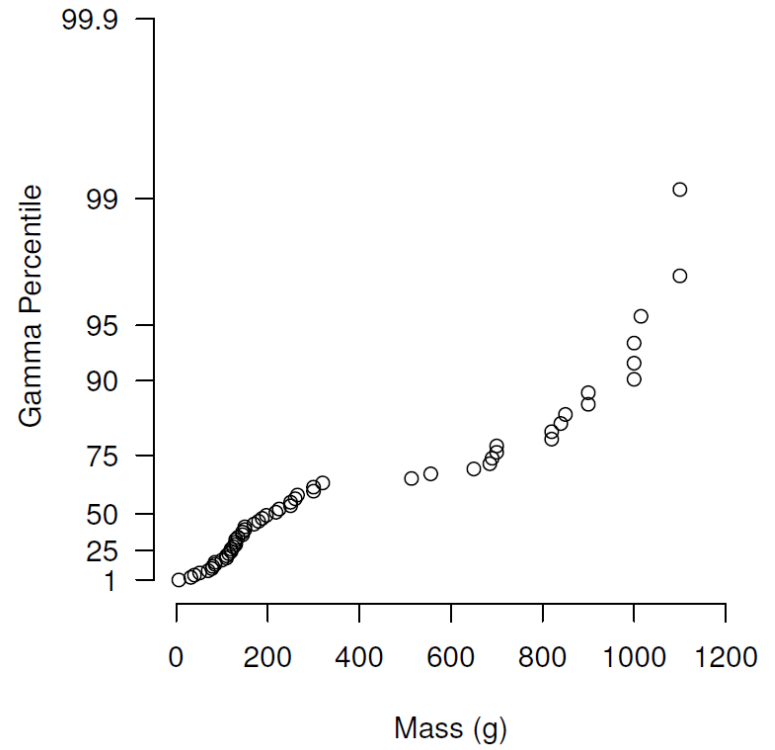
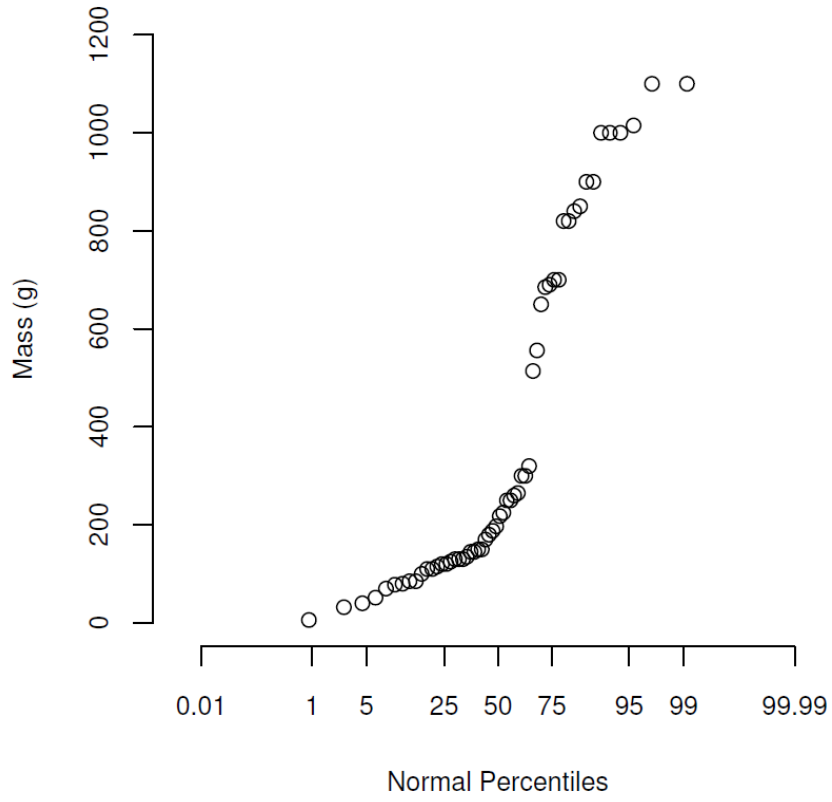
- Normal quantile plots can be used to characterize data beyond simply checking normality (cont.)
 - If there is a curved pattern with the slope increasing from left to right, then the data are skewed to the right.
 - If there is a curved pattern with the slope decreasing from left to right, then the data are skewed to the left.
 - if there is a step-like pattern with plateaus and gaps, then this is an indication that the data have been rounded (or truncated) or are discrete.

- Another distributions:
 - the gamma distribution; the beta distribution; the Chisquare distribution; and the lognormal



The Probability Plot

- a variation of the quantile-quantile plot
- The points plotted are $\{(Q1(p_k), Q2(p_k))\}$
- But the choice of scale for the reference distribution is chosen to be cumulative probability instead of quantile values





Selesai