Stat 312: Lecture 06
Quantile-quantile plots

Moo K. Chung
mchung@stat.wisc.edu

September 23, 2004

1. In order to compute $100(1-\alpha)\%$ confidence interval, it is required to find $z_{\alpha/2}$ that satisfies $P(Z > z_{\alpha/2}) = \alpha/2$ for given α. We will study how to find $z_{\alpha/2}$ and more. This lecture is based on Chapter 4.6.

2. The p-th *quantile* point q for random variable X is the point such that

$$F(q) = P(X \leq q) = p.$$

The textbook represent it in terms of *percentile*. Note that p-th quantile $= 100 \times p$-th percentile. So given p,

$$q = F^{-1}(p).$$

For $X \sim N(0,1)$, it is easy to find the p-th quntile using

```
> qnorm(1)
[1] Inf
> qnorm(0.5)
[1] 0
> qnorm(0)
[1] -Inf
> qnorm(0.5)
[1] 0
> qnorm(0.95)
[1] 1.644854
> qnorm(0.05)
[1] -1.644854
```

In order to find z_α, we use command qnorm$(1-\alpha)$.

3. Given n observations $x_1, \cdots, x_n$, we order them from the smallest to the largest and we have $x_{(1)}, \cdots, x_{(n)}$. The i-th smallest observation is defined as the $(i-0.5)/n$-th *sample quantile* point or $100(i-0.5)/n$ *sample percentile* point.

```
> library(Devore6)
> data(xmp01.05)
```

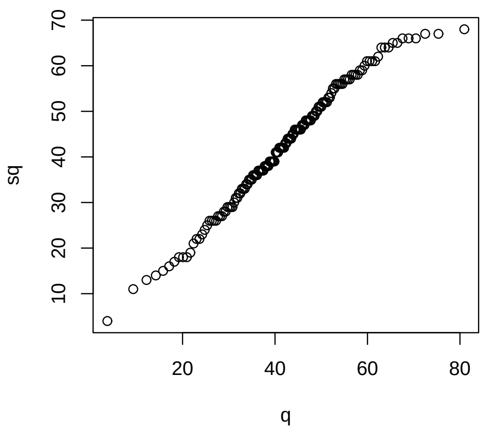

Figure 1: Plot of ordered data `bingePct` showing sample `sq`-th quantile.

```
> attach(xmp01.05)
> sq <- sort(bingPct)
```

4. If `bingePct` really follows $N(42, 14^2)$, then the sample quantiles should be resonably close to the corresponding quantiles of the normal distribution. The corresponding quantile points for `bingePct` can be computed using

```
> q=qnorm((1:140-0.5)/140,42,14)
```

We can check how closely the sample quantiles corresponds to the normal distribution by plotting the *quantile-quantile plot* (QQ-plot) of the sample quantiles vs. the corresponding quantiles of a normal distribution (Figure 1).

```
> plot(q,sq)
```

5. If $X \sim N(\mu, \sigma^2)$, then

$$Z = \frac{X-\mu}{\sigma} \sim N(0,1).$$

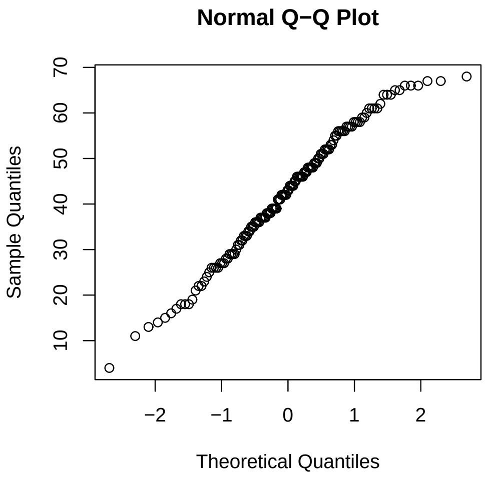

Figure 2: Normal probability plot of `bingePct`

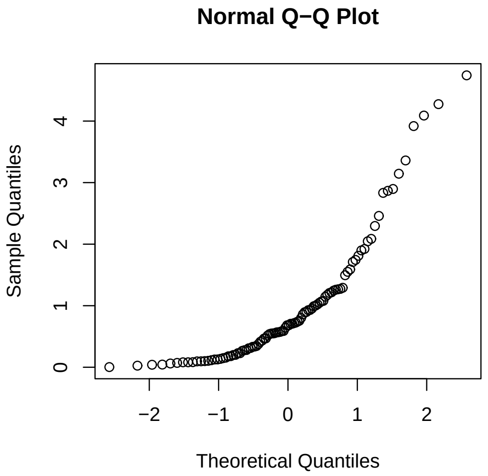

Figure 3: Normal probability plot of data that follows an exponential disribution

The p-th quantile of X is given by

$$p = P(X \leq q) = P\Big(\frac{X-\mu}{\sigma} \leq \frac{q-\mu}{\sigma}\Big) = P\Big(Z \leq \frac{q-\mu}{\sigma}\Big).$$

The quantile point q of $N(\mu, \sigma^2)$ corresponds to the quantile point $(q-\mu)/\sigma$ of $N(0,1)$. So we do not need to find the quantile points of $N(\mu, \sigma^2)$. All we need is the quantile points of $N(0,1)$ for checking normality. It can be easily done using command

```
> qqnorm(bingePct)
```

This plot (Figure 2) is usually referred as the normal probability plot.

6. What would happen if we plot data that do not follow a normal distribution?

```
> x<-rexp(100,1)
> qqnorm(x)
```

Review Problems. Example 4.28. 4.29.