

2.2 Visualizing single random variables (その 4 : pp. 30-32)

箱ひげ図を描く

boxplot (exp(lexdec\$RT))

boxplot (lexdec\$RT)

箱ひげ図の読み方

- ・ 箱は 25%地点から 75%地点
- ・ ひげは最長で箱のレンジの 1.5 倍
- ・ 以上に含まれない値は個別にプロットされ、外れ値の可能性がある
- ・ 箱の中の線は中央値
- ・ 個別の点が上側に多くあるということは、歪んだ (= symmetric ではない) データを扱っているということ

対数変換について

Figure 2.4 の上と下 (対数変換されていないものとされたもの) を比較すると、歪みが緩和されていることがわかる

これは箱ひげ図ではっきりとわかる。上と下を比べると、対数変換した下側の図の方が全体的に中央に寄っている。

対数変換するのはこのように歪みを緩和するため。そうしないと一部の外れ値が結果を大きく変えてしまう。

従って本書では対数変換を行ってから統計処理を行っている箇所が多くある

逆数による変換

対数変換ではなく、逆数を用いるとうまくいく場合もある

RT の逆数だと $1/RT$ だが、 $-1000/RT$ と -1000 倍するときもある

- ・ 1000 倍するのは小さな数を避けるため
- ・ マイナスにするのは元データで大きな値を変換後も大きな値に保つため