

Rob J Hyndman  
George Athanasopoulos

# FORECASTING

## PRINCIPLES AND PRACTICE

A comprehensive introduction to the latest forecasting methods using R. Learn to improve your forecast accuracy using dozens of real data examples.



3RD EDITION

 **OTexts**  
Oxford Texts in Finance and Statistics

## 5. The forecaster's toolbox

### 5.6 Forecasting using transformations

[OTexts.org/fpp3/](http://OTexts.org/fpp3/)

# Mathematical transformations

If the data show different variation at different levels of the series, then a transformation can be useful.

Denote original observations as  $y_1, \dots, y_n$  and transformed observations as  $w_1, \dots, w_n$ .

## Box-Cox transformations

$$w_t = \begin{cases} \log(y_t), & \lambda = 0; \\ (\text{sign}(y_t)|y_t|^\lambda - 1)/\lambda, & \lambda \neq 0. \end{cases}$$

- Natural logarithm, particularly useful because they are more interpretable: changes in a log value are **relative (percent) changes on the original scale**.

# Back-transformation

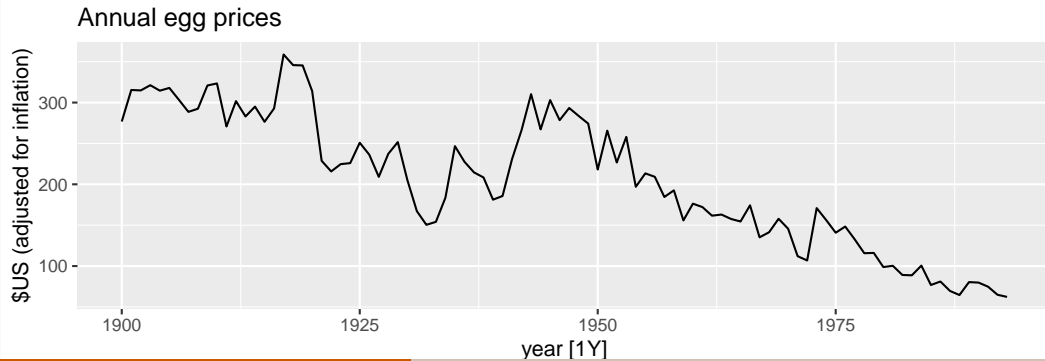
We must **reverse the transformation** or **back-transform** to obtain forecasts on the original scale. The reverse Box-Cox transformations are given by

## Box-Cox back-transformations

$$y_t = \begin{cases} \exp(w_t), & \lambda = 0; \\ \text{sign}(\lambda w_t + 1) |\lambda w_t + 1|^{1/\lambda}, & \lambda \neq 0. \end{cases}$$

# Modelling with transformations

```
eggs <- prices |>  
  filter(!is.na(eggs)) |>  
  select(eggs)  
eggs |> autoplot() +  
  labs(title = "Annual egg prices", y = "$US (adjusted for inflation)")
```



# Modelling with transformations

Transformations used in the left of the formula will be automatically back-transformed. To model log-transformed egg prices, you could use:

```
fit <- eggs |>  
  model(RW(log(eggs) ~ drift()))  
fit
```

```
## # A mable: 1 x 1  
##   `RW(log(eggs) ~ drift())`  
##                               <model>  
## 1                             <RW w/ drift>
```

# Forecasting with transformations

```
fc <- fit |>
  forecast(h = 50)
fc
```

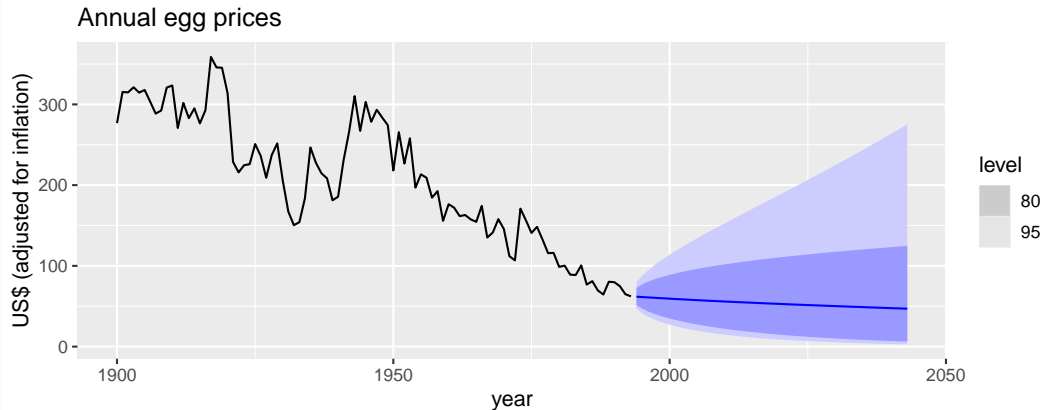
```
## # A tibble: 50 x 4 [1Y]
```

```
## # Key:           .model [1]
```

##	.model	year	eggs	.mean
##	<chr>	<dbl>	<dist>	<dbl>
##	1 RW(log(eggs) ~ drift())	1994	t(N(4.1, 0.018))	61.8
##	2 RW(log(eggs) ~ drift())	1995	t(N(4.1, 0.036))	61.4
##	3 RW(log(eggs) ~ drift())	1996	t(N(4.1, 0.055))	61.0
##	4 RW(log(eggs) ~ drift())	1997	t(N(4.1, 0.074))	60.6
##	5 RW(log(eggs) ~ drift())	1998	t(N(4.1, 0.093))	60.2
##	6 RW(log(eggs) ~ drift())	1999	t(N(4, 0.11))	59.8
##	7 RW(log(eggs) ~ drift())	2000	t(N(4, 0.13))	59.4
##	8 RW(log(eggs) ~ drift())	2001	t(N(4, 0.15))	59.0

# Forecasting with transformations

```
fc |> autoplot(eggs) +  
  labs(title = "Annual egg prices",  
        y = "US$ (adjusted for inflation)")
```



# Bias adjustment

- Back-transformed point forecasts are medians.
- Back-transformed PI have the correct coverage.



# Bias adjustment

- Back-transformed point forecasts are medians.
- Back-transformed PI have the correct coverage.

## Back-transformed means

Let  $X$  be have mean  $\mu$  and variance  $\sigma^2$ .

Let  $f(x)$  be back-transformation function, and  $Y = f(X)$ .

Taylor series expansion about  $\mu$ :

$$Y = f(X) \approx f(\mu) + (X - \mu)f'(\mu) + \frac{1}{2}(X - \mu)^2f''(\mu).$$

# Bias adjustment

- Back-transformed point forecasts are medians.
- Back-transformed PI have the correct coverage.

## Back-transformed means

Let  $X$  be have mean  $\mu$  and variance  $\sigma^2$ .

Let  $f(x)$  be back-transformation function, and  $Y = f(X)$ .

Taylor series expansion about  $\mu$ :

$$Y = f(X) \approx f(\mu) + (X - \mu)f'(\mu) + \frac{1}{2}(X - \mu)^2f''(\mu).$$

$$E[Y] = E[f(X)] \approx f(\mu) + \frac{1}{2}\sigma^2f''(\mu)$$

# Bias adjustment

**Box-Cox back-transformation:**

$$y_t = \begin{cases} \exp(w_t) & \lambda = 0; \\ (\lambda W_t + 1)^{1/\lambda} & \lambda \neq 0. \end{cases}$$

$$f(x) = \begin{cases} e^x & \lambda = 0; \\ (\lambda x + 1)^{1/\lambda} & \lambda \neq 0. \end{cases}$$

$$f''(x) = \begin{cases} e^x & \lambda = 0; \\ (1 - \lambda)(\lambda x + 1)^{1/\lambda - 2} & \lambda \neq 0. \end{cases}$$

# Bias adjustment

**Box-Cox back-transformation:**

$$y_t = \begin{cases} \exp(w_t) & \lambda = 0; \\ (\lambda W_t + 1)^{1/\lambda} & \lambda \neq 0. \end{cases}$$

$$f(x) = \begin{cases} e^x & \lambda = 0; \\ (\lambda x + 1)^{1/\lambda} & \lambda \neq 0. \end{cases}$$

$$f''(x) = \begin{cases} e^x & \lambda = 0; \\ (1 - \lambda)(\lambda x + 1)^{1/\lambda - 2} & \lambda \neq 0. \end{cases}$$

$$E[Y] \approx \begin{cases} e^\mu \left[ 1 + \frac{\sigma^2}{2} \right] & \lambda = 0; \\ (\lambda \mu + 1)^{1/\lambda} \left[ 1 + \frac{\sigma^2(1-\lambda)}{2(\lambda \mu + 1)^2} \right] & \lambda \neq 0. \end{cases}$$

# Bias adjustment

```
fc |>  
  autoplot(eggs, level = 80, point_forecast = lst(mean, median)) +  
  labs(title = "Annual egg prices",  
        y = "US$ (adjusted for inflation)")
```

