

# Abschätzung Multiple Testing Bias und Data Snooping Bias mit Landau Symbolen beim Trading

nick.albrecht

January 2026

## 1 Introduction

[a4paper,11pt]article

amsmath,amssymb,amsthm booktabs geometry margin=2.5cm

Data-Snooping- und Multiple-Testing-Bias im algorithmischen Trading

## 2 Problemstellung und Annahmen

Es wird die Verzerrung der Performance-Schätzung im algorithmischen Trading analysiert, die durch die Auswahl der besten Strategie aus einer Vielzahl getesteter Varianten entsteht. Der Fokus liegt auf dem Data-Snooping-Bias sowie auf der Kombination aus Multiple-Testing- und Data-Snooping-Effekten.

Es gelten folgende vereinfachte Annahmen:

- Beobachtungszeitraum: ein Jahr täglicher Renditen,  $T = 252$ .
- Es existiert kein echter ökonomischer Edge:

$$SR_{true} = 0.$$

- Für jede getestete Strategie  $k = 1, \dots, K$  ist der geschätzte Sharpe-Ratio näherungsweise normalverteilt:

$$\widehat{SR}_k \sim \mathcal{N}\left(0, \frac{1}{T}\right).$$

- Alle getesteten Strategien werden als unabhängig betrachtet (keine Korrektur für Korrelationen).

Diese Annahmen stellen eine idealisierte, analytisch tractable Umgebung dar und liefern konservative, gut interpretierbare Näherungen.

### 3 Data-Snooping-Bias durch Selektion der besten Strategie

Aus den  $K$  getesteten Strategien wird diejenige mit dem höchsten geschätzten Sharpe-Ratio ausgewählt:

$$\widehat{SR}_{\max} = \max_{1 \leq k \leq K} \widehat{SR}_k.$$

Unter den obigen Annahmen ergibt sich für den Erwartungswert dieses Extremwertes näherungsweise:

$$E[\widehat{SR}_{\max}] \approx \frac{\sqrt{2 \ln K}}{\sqrt{T}}.$$

Dieser Ausdruck beschreibt den erwarteten Data-Snooping-Bias im Sharpe-Ratio, der allein durch die Auswahl der besten Strategie entsteht, obwohl kein echter Edge existiert.

#### 3.1 Numerische Abschätzung für ein Jahr Daten

Für  $T = 252$  ergeben sich die in Tabelle 1 dargestellten Näherungswerte.

| Anzahl Tests $K$ | $E[\widehat{SR}_{\max}]$ |
|------------------|--------------------------|
| 200              | 0,205                    |
| 500              | 0,222                    |
| 2 000            | 0,246                    |
| 10 000           | 0,270                    |
| 100 000          | 0,302                    |
| 1 000 000        | 0,331                    |

Table 1: Erwarteter Sharpe-Bias durch Data Snooping bei  $T = 252$ .

### 4 Bias auf die erwartete Rendite

Der Sharpe-Ratio ist definiert als

$$SR = \frac{\mu}{\sigma},$$

wobei  $\mu$  die erwartete annualisierte Rendite und  $\sigma$  die annualisierte Volatilität bezeichnet. Daraus folgt für den Rendite-Bias:

$$\Delta\mu \approx \sigma \cdot E[\widehat{SR}_{\max}] \approx \sigma \cdot \frac{\sqrt{2 \ln K}}{\sqrt{T}}.$$

Für eine typische Aktienmarktvolatilität von  $\sigma = 20\%$  p.a. ergeben sich die in Tabelle 2 dargestellten Verzerrungen.

| $K$       | Sharpe-Bias | Rendite-Bias $\Delta\mu$ (p.a.) |
|-----------|-------------|---------------------------------|
| 200       | 0,205       | 4,10%                           |
| 500       | 0,222       | 4,44%                           |
| 2 000     | 0,246       | 4,91%                           |
| 10 000    | 0,270       | 5,41%                           |
| 100 000   | 0,302       | 6,05%                           |
| 1 000 000 | 0,331       | 6,62%                           |

Table 2: Erwarteter Rendite-Bias bei  $\sigma = 20\%$  p.a.

## 5 Kombination aus Multiple-Testing- und Data-Snooping-Bias

Beim Multiple Testing werden zahlreiche Hypothesen oder Strategien getestet, wodurch Extremwerte der Schätzstatistiken wahrscheinlicher werden. Data Snooping entsteht, wenn genau diese Extremwerte selektiert und berichtet oder implementiert werden.

Im Fall, dass alle getesteten Strategien keinen echten Edge besitzen (reine Nullhypothese), fällt der kombinierte Effekt mit dem Extremwert-Bias zusammen:

$$E[Bias_{MT \times DS}] \approx \frac{\sqrt{2 \ln K}}{\sqrt{T}}.$$

Im gemischten Fall, in dem nur ein Anteil  $q \in (0, 1]$  der getesteten Strategien keinen Edge besitzt, dominiert der Bias der Null-Strategien. Eine konservative Näherung ergibt:

$$E[Bias] \approx \frac{\sqrt{2 \ln(qK)}}{\sqrt{T}},$$

sowie für die Rendite:

$$\Delta\mu_{bias} \approx \sigma \cdot \frac{\sqrt{2 \ln(qK)}}{\sqrt{T}}.$$

Für große Suchräume ist typischerweise  $q \approx 1$ , sodass der kombinierte Bias nahezu identisch mit dem reinen Data-Snooping-Bias ist.

## 6 Asymptotische Wachstumsordnung

Aus den obigen Ausdrücken folgt für den Bias im Sharpe-Ratio die asymptotische Ordnung:

$$E[\widehat{SR}_{\max}] = \Theta\left(\sqrt{\frac{\ln K}{T}}\right).$$

Für den Bias der erwarteten Rendite gilt entsprechend:

$$\Delta\mu = \Theta\left(\sigma \sqrt{\frac{\ln K}{T}}\right).$$

Damit ergibt sich folgende Wachstums-Hierarchie:

- logarithmisches Wachstum in der Anzahl getesteter Strategien  $K$ ,
- inverse Wurzel-Abhängigkeit von der Stichprobengröße  $T$ ,
- lineare Skalierung mit der Volatilität  $\sigma$ .

## 7 Schlussbemerkung

Die Analyse zeigt, dass bereits bei einem moderaten Backtest-Horizont von einem Jahr und einer großen Anzahl getesteter Strategien erhebliche Verzerrungen der geschätzten Performance auftreten können. Insbesondere im Zusammenspiel von Multiple Testing und Data Snooping kann der erwartete Bias die Größenordnung realistischer ökonomischer Edges erreichen oder übersteigen, was die Notwendigkeit strikter Validierungs- und Korrekturverfahren im quantitativen Trading unterstreicht.