Notes: Time Series Maths

# Time Series Forecasting

## Lag:

A screenshot of a computer program

AI-generated content may be incorrect.

## Rolling Statistics:

A black screen with white text

AI-generated content may be incorrect.

## Calendar encodings:

A black screen with white text

AI-generated content may be incorrect.

# Logits and Coherence:

Here’s the crisp idea:

* **Logits → Softmax → Shares:**  
  The layer Dense(1) outputs arbitrary real numbers per item (the **logits**). Passing them through Softmax(axis=-1) gives **shares**

si=ezi∑j=1Iezjs\_i = \frac{e^{z\_i}}{\sum\_{j=1}^I e^{z\_j}}si​=∑j=1I​ezj​ezi​​

with si≥0s\_i \ge 0si​≥0 and ∑isi=1\sum\_i s\_i = 1∑i​si​=1. Logits themselves don’t ensure anything; the **softmax** does: it turns logits into a simplex (probability-like) vector that sums to 1.

* **Shares × Brand scalar = Coherent item forecasts:**  
  With brand forecast y^B(t)\hat{y}\_B(t)y^​B​(t), the item forecasts are

y^i(t)=si(t) y^B(t).\hat{y}\_i(t) = s\_i(t)\,\hat{y}\_B(t).y^​i​(t)=si​(t)y^​B​(t).

Summing over items: ∑iy^i(t)=y^B(t)∑isi(t)=y^B(t)⋅1=y^B(t)\sum\_i \hat{y}\_i(t) = \hat{y}\_B(t)\sum\_i s\_i(t) = \hat{y}\_B(t)\cdot 1 = \hat{y}\_B(t)∑i​y^​i​(t)=y^​B​(t)∑i​si​(t)=y^​B​(t)⋅1=y^​B​(t).  
That’s coherence **by construction**.

* **Tiny numeric example (3 items):**  
  Suppose at time ttt the Dense layer emits logits z=[1.0,  0.0,  −1.0]z=[1.0,\;0.0,\;-1.0]z=[1.0,0.0,−1.0] and the brand forecast is y^B(t)=200\hat{y}\_B(t)=200y^​B​(t)=200.
  + Softmax shares:  
    e1=2.718,  e0=1,  e−1=0.368e^{1}=2.718,\; e^{0}=1,\; e^{-1}=0.368e1=2.718,e0=1,e−1=0.368; sum =2.718+1+0.368=4.086= 2.718+1+0.368=4.086=2.718+1+0.368=4.086.

s=[2.7184.086,  14.086,  0.3684.086]≈[0.665,  0.245,  0.090].s = \left[\tfrac{2.718}{4.086},\; \tfrac{1}{4.086},\; \tfrac{0.368}{4.086}\right] \approx [0.665,\;0.245,\;0.090].s=[4.0862.718​,4.0861​,4.0860.368​]≈[0.665,0.245,0.090].

* + Item forecasts:  
    y^1=0.665×200=133.0\hat{y}\_1=0.665\times200=133.0y^​1​=0.665×200=133.0,  
    y^2=0.245×200=49.0\hat{y}\_2=0.245\times200=49.0y^​2​=0.245×200=49.0,  
    y^3=0.090×200=18.0\hat{y}\_3=0.090\times200=18.0y^​3​=0.090×200=18.0.
  + Sum: 133.0+49.0+18.0=200.0=y^B(t)133.0+49.0+18.0=200.0=\hat{y}\_B(t)133.0+49.0+18.0=200.0=y^​B​(t). ✔️

So the **logits** provide flexible, unconstrained scores; **softmax** converts them into nonnegative shares that sum to 1; multiplying by the **brand scalar** guarantees the item forecasts **sum exactly** to the brand forecast.