CHEAT ETC3550 SHEETS

Concise notes for forecasting exams, covering key concepts, methods, and models based on "Forecasting: Principles and Practice (3rd ed)".

Fundamentals & Time Series Decomposition

Forecasting Principles

Time	Series	Com	ponent
11110	001103	00111	ponent

Time Carica Diata

orecasting Principles	Time Series Components		Time Series Plots	
Forecast Error: Difference between observed and predicted values. Residuals: Forecast errors for the training data.	Additive Decomposition: Data = Trend + Seasonal + Random	Multiplicative Decomposition: Data = Trend * Seasonal * Random	Seasonal Plot : Data are grouped by season (e.g., months or quarters) and plotted to highlight seasonal patterns, showing how the series varies within each season.	
Goals: Minimize forecast error, avoid bias, and capture patterns.	Trend : Long-term direction of the series.	Seasonal: Regular, predictable variations that recur over a fixed	Time Plot : Time series data are plotted against time, revealing trends, seasonality, and cyclical patterns over time.	
and actual values. Indicated by non-zero mean error. Forecast Distribution: Use simulations of the	Cyclic : Fluctuations around the trend, usually over a longer period than seasonality.	period. Random: Irregular, unpredictable variations.	Scatter Plot: Data points are plotted as individual points to visualize the relationship between two variables, such as the series and its lagged values helping to identify autocorrelation and patterns.	
future to develop forecast distributions Overfitting: Fitting model too closely to the			Autocorrelation Function (ACF): Measures the correlation between a time series and its lagged values, revealing the strength and significance of autocorrelation at different lags.	
data.	Classical Decomposition: Method for decomposing a time series into its components (trend, seasonal, and irregular).	STL Decomposition: Versatile and robust method for decomposing time series data, handling both additive and multiplicative seasonality, as well as complex seasonal patterns.		
Underfitting: Model is too simple to capture patterns in the data			Partial Autocorrelation Function (PACF) : Measures the correlation between a time series and its lagged values after removing the effects of intermediate lags, isolating the direct relationship between the series and each lag.	

Simple Forecasting Methods

Basic Methods

Average Method : Forecast all future values using the average of historical data.	Assumptions about residuals: Uncorrelated, mean zero, constant variance, normally distributed.	
Formula: $ht{y}_{t+h t} = bar{y} = (y_1 + y_2 + + y_T) / T$	Plots: Time series plot, histogram, ACF plot.	
Naive Method: Forecast equals the last observed value.	Tests: Ljung-Box test.	
Formula: $ht{y}_{t+h t} = y_t$	Ljung-Box Test: Tests whether a group of autocorrelations of a time series are different from zero. Q^* = T(T+2) \sum_{k=1}^h r_k^2(e)	
Seasonal Naive Method : Forecast equals the last observed value from the same season.		
Formula: $hat{y}_{t+h t} = y_{t+h-m(k+1)}$, where m is the seasonal period and k is the integer part of (h-1)/m		

Drift Method: Forecast is the last value plus an average change over time.

Formula: $hat{y}_{t+h|t} = y_t + h frac{y_T - y_1}{T-1}$

Exponential Smoothing Methods

Simple Exponential Smoothing (SES)

Suitable for: Data with no trend or seasonality.		
Formula: $hat{y}_{t+1 t} = \alpha y_t + (1 - \alpha)$	Equat	
$hat{y}_{t t-1}$	Level:	
\alpha : Smoothing constant (0 < \alpha < 1).	b_{t-	
Higher values give more weight to recent	Trenc	
observations.	\beta	
Initialization : \hat{y}_{1 0} can be set to y_1 or the average of the first few observations.		

Holt's Linear Trend Method

Suitable for: Data with a trend but no seasonality.
Equations:
Level: $\ell = \lambda + (1 - \lambda) (\ell + 1) + (1 - \lambda) + (1 - \lambda) (\ell + 1) + (1 - \lambda) + ($
b_{t-1})
Trend: $b_t = beta^* (ell_t - ell_{t-1}) + (1 - $
\beta^*) b_{t-1}
Forecast: $hat{y}_{t+h t} = ell_t + h b_t$

a: Smoothing constant for the level. **^***: Smoothing constant for the trend.

Initialization: \ell_0 and b_0 can be estimated using linear regression on the historical data.

Damped Trend Methods

Damped Trend Methods : Similar to Holt's method, but the trend is damped over time.
$\label{eq:Formula: hat{y}_{t+h t} = \label{eq:Formula: hat{y}_{t+h t} = \label{Formula: hat{y}_{t+h t} = \$
\phi : Damping parameter (0 < \phi < 1). As h increases, the forecast approaches \ell_T + \frac{\phi}{1-\phi} b_T.

Residual Diagnostics

Holt-Winters' Seasonal Method

Suitable for: Data with both trend and seasonality. Can be additive or multiplicative.

$$\label{eq:matrix} \begin{split} & \text{Multiplicative: } hat\{y\}_{t+h}t\} = (\ell_t + hb_t)s_{t+h-m(k+1)} \\ & \text{Level: } ell_t = \alpha(y_t / s_{t-m}) + (1 - \alpha) \\ & (\ell_{t-1} + b_{t-1}) \\ & \text{Trend: } b_t = \beta^*(\ell_t - \ell_{t-1}) + (1 - \beta^*)b_{t-1} \\ & \text{Seasonal: } s_t = \gamma(y_t / (\ell_{t-1}) + b_{t-1}) \\ & \text{How the seasonal: } s_t = \alpha \\ & \text{How the seasonal: } s_t = \alpha \\ & \text{How the seasonal: } s_t = \beta \\ & \text{How the seasonal$$

Parameters: \alpha, \beta^*, \gamma are
smoothing constants (0 < \alpha, \beta^*,
\gamma < 1).</pre>

Initialization: Requires initial estimates for level, trend, and seasonal components.

ARIMA Models

ARIMA Model Components

AR(p): Autoregressive model of order p. Uses past values to predict future values.

MA(q): Moving average model of order q. Uses past forecast errors to predict future values.

I(d): Integrated component of order d. Represents the number of differences required to make the time series stationary.

Stationarity: A time series is stationary if its statistical properties (mean, variance, autocorrelation) do not change over time.

Differencing: Used to make a time series stationary. First difference: $y'_t = y_t - y_{t-1}$. Seasonal difference: $y'_t = y_t - y_{t-1}$ Model Selection

ACF and PACF Plots: Used to identify the order of AR and MA components.

Information Criteria: AIC, AICc, BIC. Lower values indicate better model fit.

AIC (Akaike Information Criterion): AIC = -2 $\log(L) + 2k$, where L is the likelihood and k is the number of parameters.

AICc (Corrected AIC): AICc = AIC + \frac{2k(k+1)}{T-k-1}, where T is the number of observations.

BIC (Bayesian Information Criterion): BIC = -2 \log(L) + k \log(T)

ARIMA Model Equations

AR(p) Model: y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + ... + \phi_p y_{t-p} + e_t

MA(q) Model: $y_t = c + e_t + \text{theta_1} e_{t-1} + \text{theta_2} e_{t-2} + ... + \text{theta_q} e_{t-q}$

ARIMA(p,d,q) Model: Combines AR(p), I(d), and MA(q) components. Requires differencing the series d times to achieve stationarity.