# Business, Economic and Financial Data — Exam

4 questions — allotted time: 1 hour

## 1) The problem of Autocorrelated Residuals in regression models.

Explain how to detect autocorrelation (Durbin–Watson, Breusch–Godfrey, ACF/PACF of residuals).

Describe remedial actions: Newey–West (robust SEs), GLS / Cochrane–Orcutt, modelling ARMA errors, and when each is appropriate.

## 2) Describe the Gradient Boosting algorithm.

Give the core idea: iteratively fit weak learners (usually trees) to the negative gradient (residuals) of the loss function.

Explain learning rate (shrinkage), number of trees, tree depth, subsampling and regularization to prevent overfitting.

## 3) [R code — summary] Two models (Bass and Generalized Bass with exponential shock) with R^2 reported.

Identify the models from coefficient patterns (innovation p, imitation q, shock parameter delta).

Compare dynamics and fit; discuss why R^2 > 0.95 (cumulative structure, strong explanatory fit, or potential overfitting).

| Model | R^2 | Key coefficients |
|---|---|---|
| Bass | 0.912 | p=0.021, q=0.145 |
| Gen. Bass (exp shock) | 0.973 | p=0.018, q=0.138, delta=0.42 |

## 4) [R code — coefficients] Identify the model and report the formula.

Given AR and MA coefficients plus exogenous regressors, this suggests an ARMAX(p,q,m) model.

Report general formula: y_t = c + Σ_{i=1}^p ϕ_i y_{t-i} + Σ_{j=1}^q θ_j ε_{t-j} + β' x_t + ε_t.

| Coefficient | Estimate | Std. Error |
|---|---|---|
| ar1 | 0.62 | 0.07 |
| ma1 | -0.35 | 0.09 |
| xreg_price | -0.11 | 0.02 |
| xreg_promo | 0.22 | 0.05 |

# Business, Economic and Financial Data — Exam

4 questions — allotted time: 1 hour

## 1) ARMAX: define the model and when it is suitable.

Define ARMAX: ARMA components (autoregressive + moving average) with exogenous regressors: ARMAX(p,q,m).

Suitable when the dependent series shows serial correlation and is influenced by external inputs; estimation via conditional MLE or least squares with diagnostics.

## 2) Define $R^2$. Is it good for model evaluation and why?

$R^2 = 1 - SSR/SST$. Adjusted $R^2$ adjusts for number of regressors.

Limitations: can be misleading in time series (nonstationarity, autocorrelation), and it does not measure predictive performance—prefer out-of-sample RMSE, AIC/BIC.

## 3) GGN competition between 2 products: describe parameters, model and evolution forecast.

Describe parameters (asymmetric growth, scale and rate parameters per product, interaction terms for competition).

Explain R-style output (estimates, std. errors, t-values) and interpret the 'evolution forecast' as predicted trajectories for each product, possibly under shocks.

| Param | Est. | Std. Err. | t value | p-value |
|---|---|---|---|---|
| a1 | 1200 | 45 | 26.7 | <0.001 |
| b1 | 0.031 | 0.007 | 4.4 | 0.0001 |
| a2 | 980 | 38 | 25.8 | <0.001 |
| b2 | 0.028 | 0.006 | 4.7 | <0.001 |

## 4) GAM: complete missing values (df) and interpret, how to improve (F-stat).

Explain effective degrees of freedom for smooth terms, how to read them in summary.gam output, and interpret F-statistics for smooth terms.

Improvement: adjust smoothing parameter via GCV/CV, change basis or number of knots, penalize to avoid overfitting.

# Business, Economic and Financial Data — Exam

4 questions — allotted time: 1 hour

## 1) Simple exponential smoothing: why is it called 'exponential'?

Because weights on past observations decline exponentially: $s_t = \alpha x_t + (1-\alpha) s_{t-1}$, so weight for k-lag is $\alpha(1-\alpha)^{k}$.

Discuss smoothing parameter $\alpha$ and its role.

## 2) Bias–Variance tradeoff

Define bias and variance; decompose expected squared error into $\text{bias}^2$ + variance + irreducible error.

Discuss implications for model complexity and methods to control tradeoff (regularization, model selection).

## 3) GBM: what is the step from BM to GBM

From additive Brownian Motion $dX_t = \mu\, dt + \sigma\, dW_t$ to multiplicative Geometric Brownian Motion $dS_t = \mu S_t\, dt + \sigma S_t\, dW_t$.

Explain lognormal distribution of $S_t$ and relevance to asset prices.

## 4) KNN Regression

Describe k-NN regression: predict by averaging outcomes of k nearest neighbors.

Discuss choice of k, distance metric, curse of dimensionality, and bias/variance effects.

# Business, Economic and Financial Data — Exam

4 questions — allotted time: 1 hour

## 1) Holt's exponential smoothing

Describe Holt's method with level and trend components and forecasting equations.

Contrast with simple exponential smoothing: Holt handles linear trend via additional smoothing equation.

## 2) Smoothing splines vs. regression splines

Smoothing splines: penalized least squares with smoothing parameter $\lambda$ producing a global solution.

Regression splines: piecewise polynomial with fixed knots—more manual control over knot placement.

## 3) GGM: step from BM to GGM

Explain GGM (Generalized Growth/Gompertz/Gamma depending on context) as extension adding shape/scale parameters for more flexibility over baseline model.

Describe uses and parameter interpretation.

## 4) Why collinearity is a problem for linear regression and how to test for it

Collinearity inflates variance of coefficient estimates and makes them unstable; tests: Variance Inflation Factor (VIF), condition number, eigenvalues.

Remedies: ridge/regularization, PCA, drop or combine variables.