FinQuiz Formula Sheet CFA Program Level II

Reading 4: Introduction to Regression

1. Sample Cov (X, Y) $= \frac{\sum_{i=1}^{n}(X_i-\bar{X})(Y_i-\bar{Y})}{n-1}$

2. Correlation Coefficient = $r_{XY} = \frac{cov_{XY}}{(s_X)(s_Y)}$ or
$r = \frac{cov(X,Y)}{\sqrt{var(X)}\sqrt{var(Y)}}$

3. t-test (for normally distributed variables) =
$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$ t distribution with (n − 2) deg. of freedom

4. Linear Regression = $Y_i = b_0 + b_1X_i + \varepsilon_i$,
   - Intercept (b_0) = $\overline{b_0} = \bar{y} - b_1\bar{x} =$
   - Slope or regression coefficient = $b_1 = \frac{cov(x,y)}{var(x)}$ or $= \frac{\sum(x-\bar{x})(y-\bar{y})}{\sum(x-\bar{x})^2}$

5. Standard Error of Estimate SEE = $S_E = \sqrt{\frac{SSE}{n-k-1}} = \sqrt{\frac{\sum_{i=1}^{n}(y_i-\hat{y})^2}{n-k-1}}$

6. Coefficient of Determination (R^2) =
$= \frac{SST-SSE}{SST} = \frac{RSS}{SST}$ where, $0 \leq R^2 \leq 1$
(for single independent variable $R^2 = r^2$)

7. SST = SSE + SSR(or RSS)

8. Hypothesis Testing:
   - Null and Alternative hypotheses
   - H_0: $b_1 = 0$ (no linear relationship)
   - H_1: $b_1 \neq 0$ (linear relationship does exist)
   - Test statistic = $t = \frac{\hat{b}_1 - b_1}{s_{b_1}}$
   - Confidence Interval = $b_1 \pm t_c s_{b_1}$

9. ANOVA (Analysis of variance) =

ANOVA	SS	MSS	F
Regression df = k	$SSR = \sum_{i=1}^{n}(\hat{y}_i - \bar{y})^2$	$\frac{SSR}{k}$	$\frac{SSR/k}{SSE/(n-k-1)}$
Error df = n-k-1	$SSE = \sum_{i=1}^{n}(y_i - \hat{y})^2$	$\frac{SSE}{n-k-1}$	
Total df = n-1	$SST = \sum_{i=1}^{n}(y_i - \bar{y})^2$		

Source of Variability	DoF	Sum of Squares	Mean Sum of Squares
Regression (Explained)	1	RSS	MSR = RSS/1
Error (Unexplained)	n-2	SSE	MSE = SSE/n-2
Total	n-1	SST= RSS + SSE	

10. F-Statistic or F-Test = $\frac{MSR}{MSE} = \frac{(\frac{RSS}{k})}{(\frac{SSE}{n-k-1})}$
(df numerator = k = 1)
(df denominator = n – k – 1 = n – 2)

11. Prediction Intervals = $\hat{Y} \pm t_c s_f$

where $s_f^2 = s^2\left[1+\frac{1}{n}+\frac{(X-\bar{X})^2}{(n-1)s_X^2}\right]$ and

$s_f = \sqrt{s_f^2}$

Reading 5: Multiple Regression & Issues in Regression Analysis

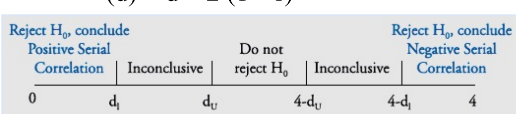
1. $Y_i = b_0 + b_1X_{1i} + b_2X_{2i} + \dots + b_kX_{ki} + \varepsilon_i, i = 1, 2, \dots n$

2. Prediction equation = $\hat{Y}_i = \hat{b}_0 + \hat{b}_1X_{1i} + \hat{b}_2X_{2i}+\dots+\hat{b}_kX_{ki} + \varepsilon_i, i$

3. Adjusted R² = $\bar{R}^2 = 1 - \left(\frac{n-1}{n-k-1}\right)(1-R^2)$

4. Breusch–Pagan test
   - H_0 = No conditional Heteroskedasticity exists
   - H_A = Conditional Heteroskedasticity exists
   - Test statistic = n × $R^2_{residuals}$

5. Durbin-Waston Test = $DW = \frac{\sum_{t=2}^{T}(\hat{\varepsilon}_t-\hat{\varepsilon}_{t-1})^2}{\sum_{t=1}^{T}\hat{\varepsilon}_t^{\,2}}$
   - For Large Sample size DW Statistic (d) = d ≈ 2 (1 – r)



Reading 6: Time Series Analysis

1. Linear Trend Models = $y_t = b_0 + b_1t + \varepsilon_t$
   - Predicted/fitted value of y_t in period (T + 1) = $\hat{y}_{t+1} = \hat{b}_0 + \hat{b}_1(T+1)$

2. Log-Linear Trend Models = $y_t = e^{b_0+b_1t}$

3. Autoregressive Time-Series Models:
   - First order autoregressive AR (1) = $x_t = b_0 + b_1 x_{t-1} + \varepsilon_t$
   - pth-order autoregressive AR (p) = $x_t = b_0 + b_1 x_{t-1} + b_2 x_{t-2} + \dots + b_p x_{t-p} + \varepsilon_t$

4. Mean reverting level of $x_t = \frac{b_0}{1-b_1}$

5. Chain Rule of Forecasting:
   - One-period ahead forecast = $\hat{x}_{t+1} = \hat{b}_0 + \hat{b}_1x_t$
   - Two-period ahead forecast= $\hat{x}_{t+2} = \hat{b}_0 + \hat{b}_1x_{t+1}$

6. Random Walks and Unit Roots:
   - Random Walk without drift = $x_t = x_{t-1} + \varepsilon_t$ where, $b_0 = 0$ and $b_1 = 1$.
   - Correcting Random Walk = $y_t = x_t - x_{t-1}$
   - Random walk with a drift = $x_t = b_0 + x_{t-1} + \varepsilon_t$ where, $b_0 \neq 0$ and $b_1 = 1$
   - By taking first difference $y_t = x_t - x_{t-1} = b_0 + \varepsilon_t$

7. Using Dickey-Fuller Test = $x_t - x_{t-1} = b_0 + (b_1 - 1)\, x_{t-1} + \varepsilon_t$

8. Smoothing Past Values with n-Period Moving Average = $\frac{x_t + x_{t-1} + x_{t-2} + \dots + x_{t-(n-1)}}{n}$

9. Correcting Seasonality in Time Series Models:
   - For quarterly data = $x_t = b_0 + b_1x_{t-1} + b_2x_{t-4} + \varepsilon_t$
   - For monthly data = $x_t = b_0 + b_1x_{t-1} + b_2x_{t-12} + \varepsilon_t$

10. ARCH model = $\hat{\varepsilon}^2_t = \alpha_0 + \alpha_1\hat{\varepsilon}^2_{t-1} + \mu_t$ where μ_t is an error term
    - Predicting variance of errors in period t+1 = $\hat{\sigma}^2_{t+1} = \hat{\alpha}_0 + \alpha_1\hat{\varepsilon}^2_t$

Reading 7: Machine Learning

LASSO:

1. Penalty term (when λ > 0) = $\lambda\sum_{k=1}^{K}|\bar{b}_k|$

2. $\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2 + \lambda\sum_{k=1}^{K}|\hat{b}_K|$

3. When λ = 0, LASSO penalized regression = OLS regression

Reading 8: Big Data Projects

1. $X_{i(normalized)} = \frac{X_i - X_{min}}{X_{max} - X_{min}}$

   where X_i = value of observation

Performance Metrics:

2. Accuracy = (TP + TN)/(TP + FP + TN + FN)
   F1 score = (2*P*R)/(P + R)

3. Receiver Operating Characteristic (ROC):
   False positive rate (FPR) = FP/(TN + FP) and
   True positive rate (TPR) = TP/(TP + FN), which is same as recall

4. Root Mean Square Error (RMSE):

$$\sum_{i=1}^{n} \frac{(Predicted_i - Actual_i)^2}{n}$$

Reading 9: Excerpt from 'Probabilistic Approaches, Scenario Analysis, Decision Tree & Simulations'

Reading 10: Currency Exchange Rates

1. Bid-offer Spread = Offer price – Bid price
2. Fwd rate = Spot Exchange rate + $\frac{\text{Forward points}}{10{,}000}$
3. Forward premium/discount (in %) = $\frac{\text{spot exchange rate} - (\text{forward points}/10{,}000)}{\text{spot exchange rate}} - 1$

4. To convert spot rate into forward quote:
   - Spot exchange rate × (1 + % premium)
   - Spot exchange rate × (1 - % discount)

5. Covered interest rate parity:
   - $(1 + i_d) = S_{f/d}(1 + i_f)\left(\frac{1}{F_{f/d}}\right)$
   - $F_{f/d} = S_{f/d}\left(\frac{1+i_f}{(1+i_d)}\right)$
   - Using day count convention:

$$\left(1 + i_d\left[\frac{Actual}{360}\right]\right) = S_{f/d}\left(1 + i_f\left[\frac{Actual}{360}\right]\right)\left(\frac{1}{F_{f/d}}\right)$$

   - $$F_{f/d} = S_{f/d}\left(\frac{1 + i_f\left[\frac{Actual}{360}\right]}{1 + i_d\left[\frac{Actual}{360}\right]}\right)$$

6. Uncovered Interest Rate Parity :
   - $i_f - \%\Delta S^e{}_{f/d} = i_d$
   - $\%\Delta S^e{}_{f/d} = i_f - i_d$
   - Forward premium or discount:
   - For one year horizon =

$$F_{f/d} - S_{f/d} = S_{f/d}\left(\frac{i_f - i_d}{1 + i_d}\right) \cong S_{f/d}(i_f - i_d)$$

   - Using day count convention:

$$F_{f/d} - S_{f/d} = S_{f/d}\left(\frac{\left[\frac{Actual}{360}\right]}{1 + i_d\left[\frac{Actual}{360}\right]}\right)(i_f - i_d)$$

7. Forward discount or premium as % of spot rate:

$$\frac{F_{f/d} - S_{f/d}}{S_{f/d}} \cong (i_f - i_d)$$

   If uncovered interest rate parity holds
   - $$= \frac{F_{f/d} - S_{f/d}}{S_{f/d}} = \%\Delta S^e_{f/d} \cong (i_f - i_d)$$

8. Purchasing Power parity (PPP)
   - $P_f = S_{f/d} \times P_d$
   - $S_{f/d} = P_f / P_d$

9. Relative version of PPP = $\%\Delta S_{f/d} = \pi_f - \pi_d$

10. Ex ante version of PPP = $\%\Delta S^e{}_{f/d} = \pi^e{}_f - \pi^e{}_d$

11. Real Exchange Rate

$$q_{f/d} = \left(\frac{S_{f/d}P_d}{P_f}\right) = S_{f/d}\left(\frac{P_d}{P_f}\right)$$

or

$$q_{f/d} = S_{f/d}\left(\frac{CPI_d}{CPI_f}\right)$$

12. Fisher effect:
   - $i_d = r_d + \pi^e{}_d$
   - $i_f = r_f + \pi^e{}_f$
   - $i_f - i_d = (r_f - r_d) + (\pi^e{}_f - \pi^e{}_d)$
   - $(r_f - r_d) = (i_f - i_d) - (\pi^e{}_f - \pi^e{}_d)$