1. (a) The parameter of interest is the mean credit card balance.

 (b) The factor associated with this experiment is the type of credit card.

 (c) There are four levels of the factor, corresponding to the four types of card.

(d) The number of degrees of freedom available for determining between sample variation is 4 - 1 = 3.

(e) The number of degrees of freedom available for determining within sample variation is (25 + 25 + 26 + 24) - 4 = 96.

(f) The number of degrees of freedom available for determining total variation is (25 + 25 + 26 + 24) - 1 = 99.

2. (a) Group means and variances for the eight categories of mutual fund are as under.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Mutual fund | AG | G | G-I | IF | I | AA | PM | B |
| Mean | 8.33 | 1.67 | 4.33 | 2.33 | 12.33 | 3.67 | 5.67 | 1.33 |
| Variance | 10.33 | 22.33 | 4.33 | 10.33 | 4.33 | 33.33 | 1.33 | 4.33 |

 The one-way ANOVA table for single observation per cell is as follows.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Source of Variation | Sum of Squares | Degrees of Freedom | Mean Square | F ratio |
| Between MF types | 297.625 | 7 | 42.518 | 3.752 |
| Within MF types | 181.333 | 16 | 11.333 |  |
| Total | 478.96 | 23 |  |  |

 Variation across MF categories can be tested by the F-ratio in the above table, 3.752. Under the null hypothesis of no significant difference of mean percentage gains across the categories, this ratio should have the F distribution with 7 and 16 degrees of freedom. The critical value of this distribution for 5% level of significance is 2.657, which is exceeded by the observed F-value. Therefore, it can be concluded that *there is statistically significant variation in the mean percentage gains across the categories*. For the record, the p-value of this test is 0.0135 (smaller than 0.05).

(b) The 0.05 upper quantile of studentized range distribution for degrees of freedom 8 and 16 is 4.896. Per group sample size is 3, estimated error variance is 11.333. Thus, the threshold for Tukey-Kramer multiple comparisons test is 9.516. The table of absolute group mean differences is shown below, with values above threshold marked in red.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | AG | G | G-I | IF | I | AA | PM | B |
| AG | 0.00 | 6.67 | 4.00 | 6.00 | 4.00 | 4.67 | 2.67 | 7.00 |
| G | 6.67 | 0.00 | 2.67 | 0.67 | 10.67 | 2.00 | 4.00 | 0.33 |
| G-I | 4.00 | 2.67 | 0.00 | 2.00 | 8.00 | 0.67 | 1.33 | 3.00 |
| IF | 6.00 | 0.67 | 2.00 | 0.00 | 10.00 | 1.33 | 3.33 | 1.00 |
| I | 4.00 | 10.67 | 8.00 | 10.00 | 0.00 | 8.67 | 6.67 | 11.00 |
| AA | 4.67 | 2.00 | 0.67 | 1.33 | 8.67 | 0.00 | 2.00 | 2.33 |
| PM | 2.67 | 4.00 | 1.33 | 3.33 | 6.67 | 2.00 | 0.00 | 4.33 |
| B | 7.00 | 0.33 | 3.00 | 1.00 | 11.00 | 2.33 | 4.33 | 0.00 |

 It is clear that the pairs of MF types with significantly different mean percentage gains are:

* Types G and I,
* Types IF and I,
* Types I and B.

Mutual Fund Type I has the highest mean percentage gain.

3. (a) The least squares fitted regression equation is

 y = 5.3730 + 0.94659x.

 (b) The expected actual account balance is 5.3730 + 0.94659(100) = 100.032.

 (c) 90% point of normal distribution = 1.282.

 Variance of the prediction error for Mr. Jones's actual account balance = 8.33188.

 Half-width of 90% prediction interval = $1.282×\sqrt{8.33188}$ = 3.699.

 Hence, a 90% interval estimate for Mr. Jones’s actual account balance is

 $100.032\pm 3.699$ = (96.333, 103.731).

 (d) 90% point of normal distribution = 1.282.

 Variance of the estimation error for average y when x is 100 = 7.33188.

 Half-width of 90% confidence interval = $1.282×\sqrt{7.33188}$ = 3.470.

 Hence, a 90% interval estimate for average y, when x is 100, is

 $100.032\pm 3.470$ = (96.562, 103.502).

 Interpretation: This interval estimate is narrower than the interval estimate of part (c) because it is meant to capture the *average* y, which has less variation than a particular y.

4. (a) The estimated parameters are $β\_{0}=69630.8, β\_{1}=90.37203, β\_{2}=-3629.5$. Thus, a regression equation to predict the selling price for residences is

 $y=69630.8+90.37203x\_{1}-3629.5x\_{2}$,

 where the $y$ is the selling price in dollars, $x\_{1}$ is the size in square feet and $x\_{2}$ is 1 for condominium and 0 for single family home.

 (b) The parameter $β\_{1}$ represents the increase in average selling price of a residence for every additional square foot of size for a particular type of house. For the given data, the estimated value of $β\_{1}$ is $90.37 per square foot.

 The parameter $β\_{2}$ represents the excess average price of a condominium over the average price of a single family home of same size. Since the estimated value of $β\_{2}$ is $-3629.5, it can be said that the average selling price of a condominium is $3629.5 less than the average selling price of a single family home of the same size.

 (c) For condominiums, $x\_{2}=1$. Therefore, an equation that describes the relationship between the selling price and the square footage of condominiums is

 $y=66001.3+90.37203x\_{1}$,

 where the $y$ is the selling price in dollars and $x\_{1}$ is the size in square feet.

 Likewise, an equation that describes the relationship between the selling price and the square footage of single family homes ($x\_{2}=0$) is

 $y=69630.8+90.37203x\_{1}$,

 where the $y$ is the selling price in dollars and $x\_{1}$ is the size in square feet.

 (d) The hypothesis that the relationship between the selling price and the square footage is different between condominiums and single-family homes, under the model given in part (a), is the hypothesis $β\_{2}\ne 0$. The corresponding null hypothesis is $β\_{2}=0$. This hypothesis is tested against the alternative $β\_{2}\ne 0$ by the t-statistic for $β\_{2}$, which happens to be -0.2284. With 17 degrees of freedom, the p-value happens to be 0.822. Since the p-value is very large, the hypothesis $β\_{2}=0$ cannot be rejected. Thus, one cannot reject the hypothesis that the relationship between the selling price and the square footage is the same for condominiums and single-family homes.

5. The original series, the trend line through a five-point moving average, the de-seasonalised series and the intermediate calculations are given the following table.

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Year | Quarter | Original series | Trend line | Seasonal MA | Seasonals | Crude seasonal index | Modified seasonal index | De-seasonalised series |
| 1997 | Q1 | 441.1 |   |   |   |   | 103.9% | 424.5 |
| 1997 | Q2 | 397.7 |   |   |   |   | 97.2% | 409.2 |
| 1997 | Q3 | 396.1 | 436.82 | 426.925 | 92.8% | 93.75% | 91.8% | 431.3 |
| 1997 | Q4 | 472.8 | 439.48 | 435.75 | 108.5% | 109.28% | 107.1% | 441.6 |
| 1998 | Q1 | 476.4 | 450.1 | 449.925 | 105.9% | 106.07% | 103.9% | 458.5 |
| 1998 | Q2 | 454.4 | 481.58 | 463.6 | 98.0% | 99.22% | 97.2% | 467.5 |
| 1998 | Q3 | 450.8 | 503.16 | 483.775 | 93.2% |   | 91.8% | 490.9 |
| 1998 | Q4 | 553.5 | 522.52 | 509.85 | 108.6% |   | 107.1% | 517.0 |
| 1999 | Q1 | 580.7 | 545.96 | 539.55 | 107.6% |   | 103.9% | 558.9 |
| 1999 | Q2 | 573.2 | 596.52 | 569.75 | 100.6% |   | 97.2% | 589.7 |
| 1999 | Q3 | 571.6 | 624.22 | 607.275 | 94.1% |   | 91.8% | 622.4 |
| 1999 | Q4 | 703.6 | 643.38 | 635.1 | 110.8% |   | 107.1% | 657.2 |
| 2000 | Q1 | 692 | 660.72 | 660.925 | 104.7% |   | 103.9% | 666.0 |
| 2000 | Q2 | 676.5 | 696.96 | 683 | 99.0% |   | 97.2% | 696.0 |
| 2000 | Q3 | 659.9 |   | 695.3 | 95% |   | 91.8% | 718.6 |
| 2000 | Q4 | 752.8 |   |   |   |   | 107.1% | 703.2 |

 The original series, the trend line and the de-seasonalised series are plotted below.

6. (a) The plot of the time series is given below.

 There appears to be a weakly rising trend, a seasonal (quarterly) component and some random fluctuations. Presence of any cyclic component is not obvious from the plot.

(b) The fitted line is

 $y=-7797.63+4x$,

 where $y$ is the number of new clients and $x$ is the time (2006 Q1 is 2006, 2006 Q2 is 2006.25, etc.). The resulting forecasts for the four quarters of 2009, the corresponding actual number of new clients, the absolute forecast error and the squared forecast error are reported below.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|   | Actual number of new clients | Forecast number of new clients | Absolute forecast error | Squared forecast error |
| Q1 | 218 | 250.92 | 32.92 | 1084.01 |
| Q2 | 241 | 253.48 | 12.48 | 155.67 |
| Q3 | 240 | 256.03 | 16.03 | 256.93 |
| Q4 | 231 | 258.58 | 27.58 | 760.74 |

 It follows that the MAD is 22.25 and the MSE is 495.19.

(c) The original data, together with the seasonal indexes and intermediate computations are shown in the following table.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Year | Quarter | new clients | Seasonal MA | Seasonals | Crude seasonal index | Modified seasonal index |
| 2006 | Q1 | 218 |  |  | 92.7% | 92.5% |
| 2006 | Q2 | 250 |  |  | 104.6% | 104.3% |
| 2006 | Q3 | 244 | 235.25 | 103.7% | 104.6% | 104.4% |
| 2006 | Q4 | 229 | 228.25 | 100.3% | 99.0% | 98.8% |
| 2007 | Q1 | 190 | 220.75 | 86.1% |  |  |
| 2007 | Q2 | 220 | 216.75 | 101.5% |  |  |
| 2007 | Q3 | 228 | 214.75 | 106.2% |  |  |
| 2007 | Q4 | 221 | 226.25 | 97.7% |  |  |
| 2008 | Q1 | 236 | 237.5 | 99.4% |  |  |
| 2008 | Q2 | 265 | 246.25 | 107.6% |  |  |
| 2008 | Q3 | 263 | 253 | 104.0% |  |  |
| 2008 | Q4 | 248 |  |  |  |  |

(d) The seasonal adjustments of part (c) leads to the seasonally adjusted data

|  |  |  |
| --- | --- | --- |
| Year | Quarter | De-seasonalised series |
| 2006 | Q1 | 235.644 |
| 2006 | Q2 | 239.639 |
| 2006 | Q3 | 233.760 |
| 2006 | Q4 | 231.821 |
| 2007 | Q1 | 205.378 |
| 2007 | Q2 | 210.882 |
| 2007 | Q3 | 218.431 |
| 2007 | Q4 | 223.722 |
| 2008 | Q1 | 255.101 |
| 2008 | Q2 | 254.017 |
| 2008 | Q3 | 251.962 |
| 2008 | Q4 | 251.055 |

 As in part (b), we fit a straight line to the adjusted data, which becomes

 $y=-16645+8.4086x$.

 The resulting forecasts of the seasonally unadjusted number of clients for the four quarters of 2009 are reported below.

|  |  |
| --- | --- |
|  Quarter | Seasonally unadjusted Forecast number of new clients |
| Q1 | 247.95 |
| Q2 | 250.05 |
| Q3 | 252.15 |
| Q4 | 254.25 |

(e) The seasonally adjusted forecasts for the four quarters of 2009, together with the corresponding actual numbers of new clients, the absolute forecast error and the squared forecast error, are reported below.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|   | Actual number of new clients | Seasonally adjusted forecast number of new clients | Absolute forecast error | Squared forecast error |
| Q1 | 218 | 229.38 | 11.38 | 129.58 |
| Q2 | 241 | 260.86 | 19.86 | 394.49 |
| Q3 | 240 | 263.20 | 23.20 | 538.17 |
| Q4 | 231 | 251.16 | 20.16 | 406.47 |

 It follows that the MAD is 18.65 and the MSE is 347.86.

(f) The forecasting technique used in part (e) leads to smaller MAD and MSE. This technique deserves to be recommended.

7. (a) $P\left(fund was winner in period 1\right)=0.527$.

 $P\left(fund was loser in period 1\right)=0.473$.

 $P\left(fund was winner in period 1\right)=0.472$.

 $P\left(fund was loser in period 1\right)=0.527$.

 All these conditional probabilities are empirical.

 $P\left(Fund is loser in period 1 and period 2\right)$

$$=P\left(fund was loser in period 1\right)×P\left(Fund is loser in period 1\right)=0.473×0.5=0.237.$$

 (b) Return from portfolio, $R=0.75RA+0.25RB$, where

 $μ\_{RA}=20\%$, $μ\_{RB}=12\%$, $σ\_{RA}^{2}=625$, $σ\_{RB}^{2}=196$, $σ\_{RA,RB}=120$.

 $μ\_{R}=0.75×20\%+0.25×12\%=18\%$.

 $ρ\_{RA,RB}=\frac{σ\_{RA,RB}}{σ\_{RA}σ\_{RB}}=\frac{120}{25×14}=0.3429$.

 The correlation matrix is $\left(\begin{matrix}1&0.3429\\0.3429&1\end{matrix}\right)$.

 Portfolio variance, $σ\_{R}^{2}=0.75^{2}σ\_{RA}^{2}+0.25^{2}σ\_{RB}^{2}+2×0.75×0.25σ\_{RA,RB}$,

 i.e., $σ\_{R}^{2}=0.75^{2}×625+0.25^{2}×196+2×0.75×0.25×120=408.81$.

 Portfolio standard deviation, $σ\_{R}=\sqrt{408.81}=20.22\%$.