Wk4

Week 4 Practice 2 (Regression Model and PIG) Question 7

Assume you work for a general home construction company “big box” retail chain. Your supervisor points out that there is a near-linear association between retail price of an item and how long on average it remains on the shelf. Correlation between price and time held in inventory is .82. The current model has a standard deviation of model error of 12.4 hours.

Explanation provided: Model error on standardized data is equal to sqrt(1-R^2). At R= .82, the standard deviation of the model error is .5724. The standard deviation of the model error needed would be .5724\*(8/12.4) = .3693. Since R = sqrt(1 – error-squared), the correlation needed is .93.

MY QUESTION: Why should the standard deviation of the model be linearly proportionate as in .5724\*(8/12.4) = .3693? Why should we assume this?

The relation between error deviation from normalised data where we can use the above quoted formula to the actual error deviation is always linear.if 0.5724. Corresponds to 12.4 what is the value for 8 is linear .

4.

Two stocks have the following expected annual returns:

Oil stock – expected return = 9% with standard deviation = 13%

IT stock – expected return = 14% with standard deviation = 25%

The Stocks prices have a small negative correlation: R = -.22.

**What is the Covariance of the two stocks?**

Hint: Use the Algebra with Gaussians Spreadsheet.

-.00715

5.

Two stocks have the following expected annual returns:

Oil stock – expected return = 9% with standard deviation = 13%

IT stock – expected return = 14% with standard deviation = 25%

The Stocks prices have a small negative correlation: R = -.22.

Assume return data for the two stocks is standardized so that each is represented as having mean 0 and standard deviation 1. Oil is plotted against IT on the (x,y) axis.

**What is the covariance?**

Hint: Use the Standardization Spreadsheet.

-.22

**Correct**

Standardization changes covariance to equal correlation. Standardization leaves correlation unchanged. See the **Standardization** Spreadsheet.

6.

Two stocks have the following expected annual returns:

Oil stock – expected return = 9% with standard deviation = 13%

IT stock – expected return = 14% with standard deviation = 25%

The Stocks prices have a small negative correlation: R = -.22.

**What is the standard deviation of a portfolio consisting of 70% Oil and 30% IT?**

Hint: Use either the Algebra with Gaussians or the Markowitz Portfolio Optimization Spreadsheet.

10.44%

7.

Two stocks have the following expected annual returns:

Oil stock – expected return = 9% with standard deviation = 13%

IT stock – expected return = 14% with standard deviation = 25%

The Stocks prices have a small negative correlation: R = -.22.

Use MS Solver and the Markowitz Portfolio Optimization Spreadsheet to Find the weighted portfolio of the two stocks with lowest volatility.

10.36%

**Correct**

The minimum-volatility-portfolio is weighted at .74 Oil and .26 IT.

11.

Customers who use online chat support can rate the help they receive from a customer support worker as a 0 (useless), a 1, 2, 3, 4, or 5 (excellent). The mean rating is 3.935, with standard deviation = 1.01.

A new support worker named Barbara has received, over her first 100 chat sessions, an average rating of 3.7. Her boss calls her in and threatens to fire her if her performance does not improve.

**Barbara replies “Its just bad luck - I’ve had more than my share of unhappy customers today.” Who is most likely right?**

Barbara’s performance has a Z score = (mean - actual) /standard deviation of the sample means = (3.7 - 3.935) / (1.01/sqrt(100)) = -2.327.

If Barbara’s performance is due to chance, the probability of observing a z-score as low, or lower, than -2.327 would be less than 1%.

Note that due to the Central Limit Theorem, even though the raw customer scores do not have a Gaussian distribution, the sample means will have a Gaussian distribution, so that use of the NormSDist function is appropriate. NormSDist(-2.327) = .0099898.

12.

Your company currently has no way to predict how long visitors will spend on the Company’s web site. All it known is the average time spent is 55 seconds, with an approximately Gaussian distribution and standard deviation of 9 seconds. It would be possible, after investing some time and money in analytics tools, to gather and analyzing information about visitors and build a linear predictive model with a standard deviation of model error of 4 seconds.

**How much would the P.I. G. of that model be?**

Hint: Use the Correlation and P.I.G. Spreadsheet

With no model, the correlation R = 0 and the “model error” is equal to the standard deviation of Y = 9 seconds. Standardized, the model error when R = 0 is equal to 1. Reducing the model error to 4 seconds is equivalent to reducing the standardized model error to 4/9 = .4444 Since Sqrt(1-R^2) = .4444, R^2 = 1- .4444^2 = .8958.