Quantitative Analysis

Quantitative Analysis

- Moments
- Probability
- Prob. distribution
- Sampling
- Hypothesis Testing
- Correlations & Copulas

Quantitative Analysis

- Moments
- Probability
- Prob. distribution
- Sampling
- Hypothesis Testing
- Correlations & Copulas

Mean

Mean: $\frac{\sum_{i=1}^{n} X_i}{n}$

Mode: Value that occurs most frequently

Median: Midpoint of data arranged in ascending/ descending order

Variance

Avg. of squared deviations from mean

Sample variance

$$s^2 = \frac{\sum_{i=1}^{n}(X_i - X_{mean})^2}{(n-1)}$$

Population variance

$$\sigma^2 = \frac{\sum_{i=1}^{N}(X_i - \mu)^2}{N}$$

Standard deviation = √Variance

$Var(ax+by)=a^2Var(x)+b^2Var(y)+2abCov(x,y)$

Q.
σ^2 of return of stock P= 100.0
σ^2 of return of stock Q=225.0
Cov (P,Q) =53.2
Current Holding $1 mn in P.
New Holding: shifting $ 1 million in Q and keeping
USD 3 million in stock P. What %age of risk (σ), is reduced?
Ans.
$\sigma_P=\sqrt{[w^2\sigma_A^2 + (1-w)^2 \sigma_B^2 +2w(1-w)Cov(A,B)]}$
w= 0.75
$c^2 = 100*(0.75)^2 + 225*(0.25)^2 +2*0.25*0.75*53.2$
σ_P= 9.5 old σ = √100 = 10
Reduction = 5%

Skewness

- Positively: mean> median> mode
- Negatively: mean< median< mode
- Skewness of Normal = 0

Kurtosis

- Leptokurtic: More peaked than normal (fat tails); kurtosis>3
- Platykurtic: Flatter than a normal; kurtosis<3
- Kurtosis of Normal = 3
- Excess Kurtosis = Kurtosis - 3

Q.
If distributions of returns from financial instruments are leptokurtotic. How does it compare with a normal distribution of the same mean and variance?
Ans.
Leptokurtic refers to a distribution with fatter tails than the normal, which implies greater kurtosis

Quantitative Analysis

- Moments
- Probability
- Prob. distribution
- Sampling
- Hypothesis Testing
- Correlations & Copulas

Probability

Properties

- P(A) = # of fav. Events/ # of Total Events
- $0 < P(A) < 1$, $P(A^c)=1-P(A)$
- $P(A \cup B)=P(A)+P(B)-P(A \cap B)$
- $=P(A)+P(B)$ If A,B mutually exclusive
- $P(A|B)= P(A \cap B)/P(B)$
- $P(A \cap B)=P(A|B)P(B)$
- $P(A \cap B)=P(A)*P(B)$ If A,B independent

Counting principles

- No. of ways to select r out of n objects: ${}^nC_r = n!/[r!*(n-r)!]$
- No. of ways to arrange r objects in n places: ${}^nP_r = n!/(n-r)!$

Sum rule and Bayes' Theorem

$$P(B) = P(A \cap B) + P(A^c \cap B)$$

$$P(B) = P(B/A) * P(A) + P(B/A^c) * P(A^c)$$

$$P(B/A) = \frac{P(A/B)*P(B)}{P(A/B)*P(B) + P(A/B^c)*P(B^c)}$$

Q.
The subsidiary will default if the parent defaults, but the parent will not necessarily default if the subsidiary defaults. Calculate Prob. of a subsidiary & parent both defaulting. Parent has a PD =.5% subsidiary has PD of.9%
Ans.
$P(P \cap S) = P(S/P)*P(P) = 1*0.5 = 0.5\%$

Q.
ABC was inc. on Jan 1, 2004. Its expected annual default rate of 10%. Assume a constant quarterly default rate. What is the probability that ABC will not have defaulted by April 1, 2004?
Ans.
P(No Default Year) = P(No default in all Quarters)
= (1-PDQ1)*(1-PDQ2)*(1-PDQ3)*(1-PDQ4)
PDQ1=PDQ2=PDQ3=PDQ4=PDQ
P(No Def Year) = $(1-PDQ)^4$
P(No Def Quarter) = $(0.9)^4$ = 97.4%

Prob. distribution

Discrete

Binomial

Only 2 possible outcomes: failure or success.
$P(x)={}^nC_x * p^x * (1-p)^{n-x}$

Binomial Random Variable
$E(X)=n*p$
$Var(X)=n*p*(1-p)=n*p*q$

Q.
A portfolio consists of 17 uncorrelated bonds. The 1-year marginal default prob. of each bond is 5.93%. If spread of default prob. is even over the year, Calculate prob. of exactly 2 bonds defaulting in first month?
Ans.
1-month default rate =1- (1-0.593)1/12
= 0.00508
Ways to select 2 bonds out of 17
= ${}^{17}C_2$ = 17*16/2
P(Exactly 2 defaults)
= (17*16/2)*(0.00508)2*(1-0.00508)15
= 0.325%

Poisson

Fix the expectation λ=np.
$P(x)=\lambda^x e^{-\lambda}/x!$ if x>=0
P(x)=0 otherwise

Q.
The number of false fire alarms in a suburb of Houston averages 2.1 per day. What is the (apprximate) probability that there would be 4 false alarms on 1 day?
Ans.
$P(X=x) = (\lambda^x e^{-x})/x!$
X= 2.1, x = 4
P(2.1) = 0.1

Continuous

AB

Quantitative Analysis

- Moments
- Probability
- Prob. distribution
  - Discrete
  - Continuous
    - AB
- Sampling
- Hypothesis Testing
- Correlations & Copulas

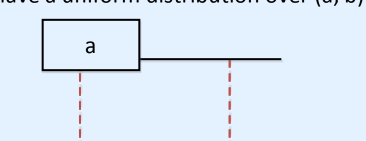
Continuous uniform distribution

- Outcome only between [a, b]
- P(outside a & b) = 0

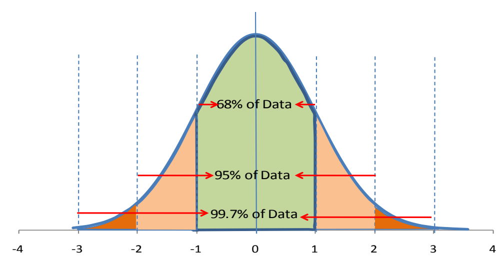
Cumulative density function (cdf) for Uniform distribution:

F(x)=0	For x <=a
F(x)=(x-a)/(b-a)	For a<x<b
F(x)=1	For x >=b

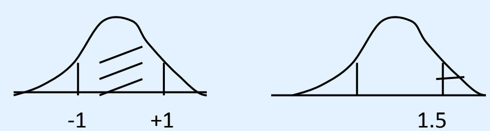
Q.
The R.V. X with density function f(X) = 1 / (b - a) for a < x < b, and 0 otherwise, is said to have a uniform distribution over (a, b). Calculate its mean.



Ans.
Since the distribution is uniform, the mean is the center of the distribution, which is the average of a and b = (a+b)/2



Q.
If Z is a standard normal R.V. An event X is defined to happen if either -1< Z < 1 or Z > 1.5. What is the prob. of event X happening if N (1) =0.8413, N (0.5) = 0.6915 and N (-1.5) = 0.0668, where N is the CDF of a standard normal variable.



Ans.
The sum of areas shown in two figures
Area 1 = 1-2*(1- N(1)) = 1-2*(0.1587)
Area 2 = 0.0668, Total Area = 0.7514

Normal Distribution (ND)

Standardized RV is normalized mean = 0, σ = 1

Z-score: # of σ a given observation is from population mean. Z=(x-µ)/σ

Q.
At a particular time, the market value of assets of the firm is \$100 Mn and the market value of debt is \$80 Mn. The standard deviation of assets is \$ 10 Mn. What is the distance to default?
Ans.
$z = (A-K)/\sigma_A = (100-80)/10 = 2$

Quantitative Analysis

Moments

Probability

Prob. distribution

Sampling

Hypothesis Testing

Correlations & Copulas

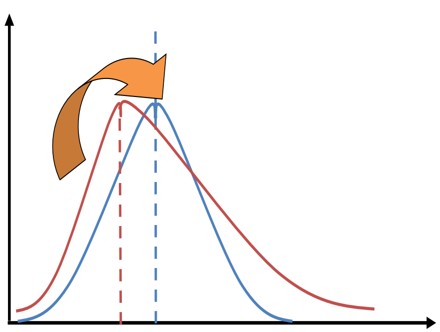
Central limit theorem

SE (σ_x) of the sample mean is σ of the dist. of sample means

- Known pop. Var. $\sigma_x = \sigma / \sqrt{(n)}$
- Unknown pop. Var. $s_x = s / \sqrt{(n)}$

Q.
25 observation are taken from a sample of known variance. Sample mean =70 and population σ = 60. You wish to conduct a two - tailed test of null hypothesis that the mean is equal to 50. What is most appropriate test statistic?
Ans.
Standard Error of mean (σ_x) = $\sigma/\sqrt{(n)} = 60/\sqrt{25} = 12$
Degrees of freedom = 24
Use t- statistic = $(x - \mu)/ \sigma_x = (70 - 50)/12 = 1.67$

As Sample Size increases Sampling Distribution Becomes Almost Normal regardless of shape of population

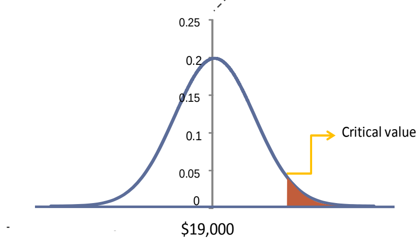


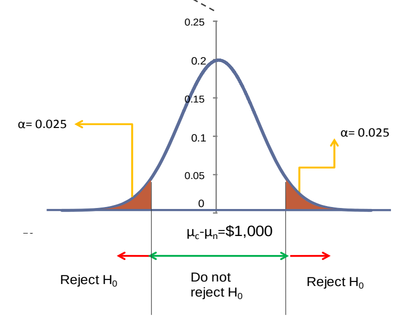
Quantitative Analysis

- Moments
- Probability
- Prob. distribution
- Sampling
- Hypothesis Testing
- Correlations & Copulas

Hypothesis Testing

- **Null hypothesis: H_0**
  - Hypothesis that the researcher wants to reject
  - Actually tested Hypothesis
  - Type I error: rejection of H_0 when it is actually true
  - Type II error: Fail to reject H_0 when it is actually false
- **Alternative Hypothesis: H_a**
  - Concluded if there is significant evidence to reject H_0
- **One tailed test**
  - Test if the value is greater than or less than K H_0: $\mu <= K$ vs. H_a: $\mu > K$
- **Two tailed test**
  - Test if the value is different from K H_0: $\mu = 0$ vs. H_a: $\mu \neq 0$
- **Z & t test**
  - If n <30 and unknown σ, use t -Test
- **P- value**
  - Given H_0 true, Prob. of obtaining value of test statistic at least as extreme as the one that was actually observed
- **2 Mean Test**
  - H_0: $\mu_1 = \mu_2$ vs H_a: $\mu_1 \neq \mu_2$
  - Use following t-statistic for unequal variances





$$t = \frac{(x_1 - x_2) - (\mu_1 - \mu_2)}{\sqrt{(s_1^2 / n_1) + (s_2^2 / n_2)}}$$

Inference Based on Sample Data	Real State of Affairs	
	H_0 is True	**H_0 is False**
H_0 is True	Correct decision Confidence level = 1- α	Type II error P (Type II error) = β
H_0 is False	Type I error Significance level = α*	Correct decision Power = 1-β
*Term α represents the maximum probability of committing a Type I error		

Quantitative Analysis

- Moments
- Probability
- Prob. distribution
- Sampling
- Hypothesis Testing
  - Null hypothesis: H_0
  - Alternative Hypothesis: H_a
  - One tailed test
  - Two tailed test
  - Z & t test
  - P- value
  - 2 Mean Test
- Correlations & Copulas

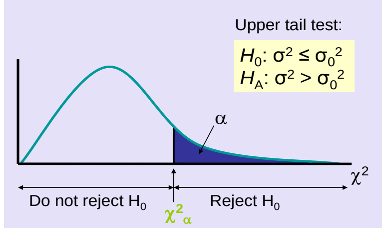
Q.
A stock has initial price of $100. It price one year from now is given by S = 100 *exp(r), where the rate of return r is normally distributed with mean of 0.1 and a standard deviation of 0.2. What is the range of S in an year with 95% confidence?
Ans.
$100e^{(0.1-1.96*0.2)} < S < 100e^{(0.1+1.96*0.2)}$
$74.68 < S < 163.56$

Q.
If standard deviation of a normal population is known to be 10 and the mean is hypothesized to be 8. Suppose a sample size of 100 is considered. What is the range of sample means in which hypothesis can be accepted at significance level of 0.05?
Ans.
$s_x = \sigma/\sqrt{n} = 10/\sqrt{100} = 1$
$z = (x-\mu)/\sigma_x$
$= (x-8)/1$

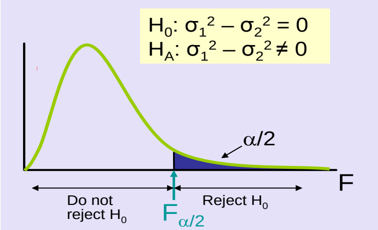
At 95% $-1.96<z<1.96$; So $6.04<x<9.96$

Hypothesis Tests for Variances

- Tests for a Single Population Variances
  - Chi-Square test
  - H_0: $\sigma^2 = c$
    H_a: $\sigma^2 \neq c$
    $$\chi^2 = \frac{(n-1)s^2}{\sigma^2}$$



- Tests for a two Population Variances
  - F test
  - H_0: $\sigma_1^2 - \sigma_2^2 = 0$
    H_A: $\sigma_1^2 - \sigma_2^2 \neq 0$
    $$F = \frac{s_1^2}{s_2^2}$$



Quantitative Analysis

Moments

Probability

Prob. distribution

Sampling

Hypothesis Testing

Correlations & Copulas

$$\rho = \frac{cov(V_1, V_2)}{SD\ (V_1)\ SD\ (V_2)}$$

Measuring Correlation:

1. **EWMA (Exponentially weighted moving average)**

   = $cov_n = \lambda cov_{n-1} + (1-\lambda)x_{n-1}y_{n-1}$

2. **GARCH**

   = $cov_n = \omega + \alpha x_{n-1}y_{n-1} + \beta cov_{n-1}$

Quantitative Analysis

- Correlation & Regression
- Volatility Estimation
- Simulation Modeling
- Modeling and Forecasting Trend
- Characterizing Cycles
- Modeling Cycles: MA, AR and ARMA Models

Quantitative Analysis

Correlation & Regression

Volatility Estimation

Simulation Modeling

Modeling and Forecasting Trend

Characterizing Cycles

Modeling Cycles: MA, AR and ARMA Models

Simple Linear Regression

LR model: $Y_i = b_0 + b_1 X_i + E_i$
Y_i = Dependent variable, estimated value of Y_i, given value of X_i
X_i = independent variable
b_0 = intercept term; represents Y if X = 0
b_1 = slope coefficient; measures change in Y for 1 unit change in X

Regression coefficient

Appropriate Test structure:
$H_0: b_1 = 0$; $H_a: b_1 \neq 0$
Test: $t_{b1} = (\hat{b}_1 - b_1)/s_{\hat{b}1}$
Decision Rule:
reject H_0 if $t > +t_{critical}$ or if $t < -t_{critical}$

Coefficient of Determination(R^2)

%age of total var. in Y explained by X
$R^2 = (SSR / SST)$
$= 1 - (SSE / SST)$
= explained variation/ total variation

Correlation Coefficient (CC)

Only the linear correlation,
$-1 < CC < 1$,
if CC = 0, X & Y are uncorrelated
$r_{x,y} = cov(x,y)/\sigma_x \sigma_y = \sqrt{R^2}$

Residual Diagnostic

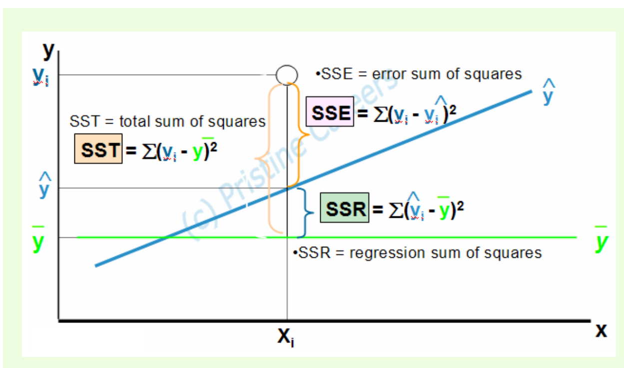
- The error variable must be normally distributed,
- The error variable must have a constant variance
- The errors must be independent of each other

Multiple Regression

$$Y_i = b_0 + b_1 X_{1i} + b_2 X_{2i} + \ldots\ldots + b_k X_{ki} + \varepsilon_i$$

Adjusted R- square is used to test the goodness of fit

$$R_a^2 = 1 - \left[\left(\frac{n-1}{n-k-1}\right) \times \left(1 - R^2\right)\right]$$



	Coefficients	Standard Error	t-statistic
Intercept	49.94	2.85	17.53
X Variable 1	-38.79	138.93	-0.28
X Variable 2	-431.75	170.50	-2.53
X Variable 3	-70.40	121.06	-0.58

Quantitative Analysis

EDU PRISTINE

- Correlation & Regression
- Volatility Estimation
- Simulation Modeling
- Modeling and Forecasting Trend
- Characterizing Cycles
- Modeling Cycles: MA, AR and ARMA Models

Volatility Estimation

EWMA

$$\sigma_n^2 = \lambda \sigma_{n-1}^2 + (1-\lambda) u_{n-1}^2$$

Where,
λ=Persistence factor/Decay Factor
1- λ= Reactive factor

Q.
Using a daily RiskMetrics EWMA model with a decay factor λ = 0.95 to develop a forecast of the conditional variance, which weight will be applied to the return that is 4 days old?

Ans.
The EWMA RiskMetrics model is defined as $h_t = \lambda * h_t - 1 + (1- \lambda) * r_t - 1$.
For t=4, and processing r_0 through the equation three times produces a factor of (1-0.95)*0.953 = 0.043 for r_0 when t = 4

GARCH

$$\sigma_n^2 = \omega + \alpha u_{n-1}^2 + \beta \sigma_{n-1}^2$$

ω =Weighted long run variance= γVL
VL=Long run avg. variance= ω/ (1-α-β)
α+β+γ=1
α+β<1 for stability so that γ is not -ve

Q.
GARCH model is estimated as follows:

$$\sigma_{t+1}^2 = 0.000005 + 0.12\mu_t^2 + 0.85\sigma_t^2$$

On a particular day 't'; actual return was -1% & the std. deviation estimate was 1.8%. Calculate the volatility estimate for next day (t+1) and long-term average volatility.

Ans.
Volatility estimate for next day
VL =.017%, Also, variance estimate for
t+1 =.000005 + 0.12*(-1%)^2 + 0.85*(1.88%)^2 = 0.0317%
Volatility (std. deviation) estimate for t+1 = sqrt(0.0317%) = 1.782%

Ans.
Long Term Volatility
In the GARCH model, 12% is the weight given to latest squared return (reactive factor). 85% is the weight given to latest variance estimate (persistence factor). Therefore, 1-0.12-0.85 = 3% is weight given to long-term average Volatility.
Therefore, 3%*VL = 0.000005 i.e. VL = 0.017%

Implied Volatility

The implied volatility of an option contract is the volatility implied by the market price of the option based on an option pricing model

Simulation Modeling

Monte Carlo Simulation

- Technique that converts uncertainties in input variables of a model into probability distributions
- Combining the distributions and randomly selecting values from them, it recalculates the simulated model many times and brings out the probability of the output

Techniques for Generating good random numbers:
- Pseudorandom number generator
- Quasirandom number generators
- Stratified sampling

Advantages of simulation modeling

- When the input random variable follows some complex distribution
- when the output is a complex function of input variable
- simulation modeling can compound probability when there are multiple input random variables
- Correlation between input variables is taken into account

Quantitative Analysis

Correlation & Regression

Volatility Estimation

Simulation Modeling

Modeling and Forecasting Trend

Characterizing Cycles

Modeling Cycles: MA, AR and ARMA Models

$$MSE = \frac{\sum_{t=1}^{T} e_t^2}{T}$$

Where T, is the sample size and $e_t = y_t - \hat{y}_t$

where, $\hat{y}_t = \widehat{\beta_0} + \widehat{\beta_1} TIME_t$

Criteria for panelizing MSE to reflect the df used:

- Akaike information criterion

$$AIC = e^{\left(\frac{2k}{T}\right)} \frac{\sum_{t=1}^{T} e_t^2}{T}$$

- Schwarz information criterion

$$SIC = T^{\left(\frac{k}{T}\right)} \frac{\sum_{t=1}^{T} e_t^2}{T}$$

- Key property for selection criteria of model is **"Consistency"**
- A model selection criteria is consistent if the following conditions are met:

1. When the true model i.e. DGP is among the models considered, the probability of selecting the true DGP approaches 1 as the sample size gets large
2. When the true model is not among the models considered, the probability of selecting the best approximation to the true DGP approaches 1 as the sample size gets large

Quantitative Analysis

Correction & Regression

Volatility Estimation

Simulation Modeling

Modeling and Forecasting Trend

Characterizing Cycles

Modeling Cycles: MA, AR and ARMA Models

- Covariance stationary

Requirements for covariance stationary are:

1. The mean of the series to be stable over time
2. The covariance structure of the series to be stable over time
3. The variance of the series – the auto covariance at displacement 0, (0) – be finite

- Autocorrelation function
- We work with autocorrelation function $p(\tau)$ rather than auto covariance function $\gamma(\tau)$

$$= \quad p(\tau) = \frac{\gamma(\tau)}{\gamma(0)}$$

White Noise:

- **When shock is uncorrelated then y_t is serially correlated and when it is not mentioned we assume this process with zero mean, constant variance and no correlation is called White Noise**
- **When y is serially independent then y is independent white noise**
- **When y is independently and identically distributed with zero mean and constant variance and uncorrelated then y is Gaussian White House**

Quantitative Analysis

- Correlation & Regression
- Volatility Estimation
- Simulation Modeling
- Modeling and Forecasting Trend
- Characterizing Cycles
- Modeling Cycles: MA, AR and ARMA Models

Moving Average Models:

- First order MA=

$$y_t = \varepsilon_t + \theta\varepsilon_{t-1} = (1+\theta L)\varepsilon_t$$

Types:

- Population autocorrelation function
- Population Partial Autocorrelation Function

Autoregressive Models (AR)

- The first order AR=

$$y_t = \varphi y_{t-1} + \varepsilon_t$$

- In lag operator form, it is

$$(1-\varphi L)y_t = \varepsilon_t$$

Types:

- Population autocorrelation function
- Population Partial Autocorrelation Function

Autoregressive Moving Average models

- MA and AR processes are combined to get better parsimonious approximations
- The ARMA (1,1) Process:

1. The process is

$$y_t = \varphi y_{t-1} + \varepsilon_t + \theta\varepsilon_{t-1}$$

2. In lag operator form, it is

$$(1-\varphi L)y_t = (1+\theta L)\varepsilon_t$$

Thank you!

Contact:

E: help@edupristine.com

Ph: +1 347 647 9001