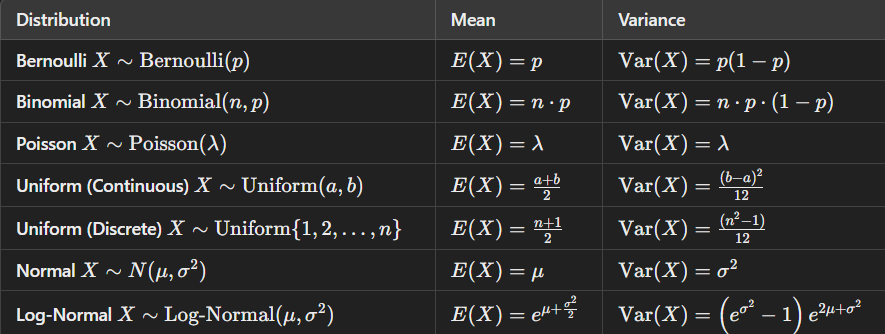
## **1. Mean of a Linear Combination** For a linear combination Z = aX + bY + c: E[Z] = a \* E[X] + b \* E[Y] + c **2. Variance of a Linear Combination** For Z = aX + bY: Var(Z) = a^2 \* Var(X) + b^2 \* Var(Y) + 2ab \* Cov(X, Y) **3.Covariance to Correlation Conversion** For random variables X and Y: ρ(X, Y) = Cov(X, Y) / (σ(X) \* σ(Y)) **Unbiasedness** Bias(θ^)=E[θ^]−θ Unbiasedness E[θ^]=θ **Consistency of Estimator** Lim P(∣θ^n​−θ∣<ϵ)=1 (n→∞​) **PDF** P(a≤X≤b)=a∫b​f(x)dx **CDF** F(x)=P(X≤x)=−∞∫x​f(t)dt **Inverse CDF** (percentile) If F(a)=pF(a) = pF(a)=p, then the inverse CDF is F−1(p)=aF^{-1}(p) = aF−1(p)=a **Probability** P(A∪B)=P(A)+P(B)-P(A∩B)​ P(A∣B)=P(A∩B)​/P(B) P(Stock Up∣Downgrade)=P(Stock Up and Downgrade)​/P(Downgrade) **Bayesian Analysis** P(A|B) = P(B|A) P(A)/P(B) **Bayesian Inference Procedure -** **Step 1: Calculate Likelihoods** - For Type A: P(3 Successes | Type A) = (0.10)^3 = 0.001 - For Type B: P(3 Successes | Type B) = (0.60)^3 = 0.216 - For Type C: P(3 Successes | Type C) = (0.80)^3 = 0.512 **Step 2: Calculate the Unconditional Probability of Observing the Data** P(3 Successes) = (0.001\*0.20) + (0.216\*0.50) + (0.512\*0.30) = 0.0002 + 0.108 + 0.1536 = 0.2618 **Step 3: Calculate Posterior Probabilities Using Bayes' Theorem** Using Bayes' theorem,: - For Type A: P(Type A | 3 Successes) = (0.001 \* 0.20) / 0.2618 = 0.000764 - For Type B: P(Type B | 3 Successes) = (0.216 \* 0.50) / 0.2618 = 0.4127 - For Type C: P(Type C | 3 Successes) = (0.512 \* 0.30) / 0.2618 = 0.586 Z is a standard normal random variable. remember that variance is chi-square distributed Normal/chi-square is t-distributed



**Linearity**: Ensures the model captures the true relationship.

**No Perfect Multicollinearity**: Avoids confusion about which variable is affecting the outcome.

**Exogeneity**: Prevents bias by ensuring no unobserved factors are influencing both predictors and the outcome.

**Homoscedasticity**: Ensures equal importance is given to all data points for accurate prediction.

**No Autocorrelation**: Ensures that each prediction is independent over time, especially for time-series data.

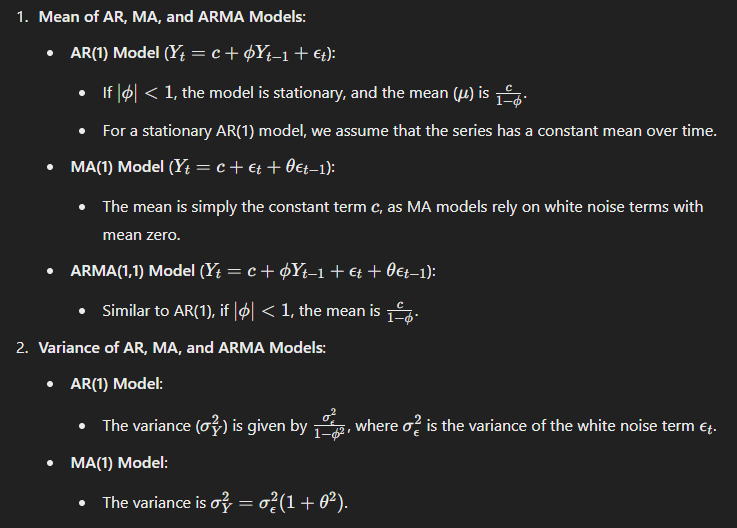
**Normality of Errors**: Allows for reliable hypothesis tests and confidence intervals.  
Violations of linearity or exogeneity can lead to **biased coefficients**.  
Violations of homoscedasticity or autocorrelation can lead to **inefficient estimates** and affect the accuracy of hypothesis tests.  
  
**OLS Conditions** -  
A1 - a 1 just says it's unbiased   
the condition for a 1 to be true is that you have a constant in your ols formula  
A5 - a 5 just says the error term is normally distributed so that you can form a T step right. The denominator of the T step needs to be Chi squared, distributed. The numerator needs to be normal. We need to assume that the error term is normally distributed, so that both the numerator and denominator would  
A2 to A4. It just says that if those are violated, then we choose a Arima-x rather than OLS

(DW) detects autocorrelation in residuals   
DW test value= 2 No issue Go for OLS   
(BG) detects autocorrelation at multiple lags   
p >0.05 Go for OLS   
(GQ) detects heteroscedasticity  
 p >0.05 Go for OLS  
The **Gauss-Markov** theorem tells us that, if we use OLS and meet its assumptions, our estimates of average height will be as close to the true value as possible compared to other methods, with the least amount of error spread.  
**R-Squared** (0-1)tells you how well these two factors (study hours and practice tests) together explain the variation in exam scores.  
**Adjusted R-Squared** (0-1) will tell you if adding a new factor, like age, actually improves the model or not. If age doesn’t help explain exam scores, adjusted R² will likely decrease.

What are the reasons for time series decomposition? You know, we just want to be a customized model that is suited to models with seasonality versus models with trend right? That's all.  
How to tell if a seasonal pattern is important or not. You can look at the amplitude of the seasonality.  
  
Arima Stationarity -  
 If it's AR(1), then the absolute value of the AR one coefficient has to be less than one. MA is always stationary.  
  
AR model: Look for sharp drop in PACF (to determine p) and gradual decay in ACF.

MA model: Look for sharp drop in ACF (to determine q) and gradual decay in PACF.

ARMA model: Look for gradual decay in both ACF and PACF, or sharp drops at the same lag if it's ARMA(1,1)



autoplot, monthdays, boxcox, nsdiffs, diff (as many times as needed), ndiffs, diff (as many times as needed), tsdisplay, Arima, check residuals  
  
VAR -  
1.Render data stationary (applies to all variables), 2. run varselect which uses BIC to select initial starting point order   
3. estimate model based on varselect   
4. run jung-box test and if pass, done, if fail, increment order and iterate until pass.   
5. if incremented order to very high number and still fail, consider differencing all the variables again

VECM -  
1.Render data stationary (applies to all variables),   
2. run varselect which uses BIC to select initial starting point order   
3. estimate model based on varselect 4. run jung-box test and if pass, done,

VECM conditions under which you can use just at least one co-integrating relationship in the data  
  
Economic intuition behind VECM -   
It's basically deviation from a long term equilibrium, which we will mean revert  
  
it's important to know that GARCH is nothing more than a special example of Arima X. So the reason why this connection is important  
is that if you want to fit a GARCH model if you want to. you have a new data set. You have an asset class that no one has seen before. Currently, it's fashionable with digital currencies. Right? So it's a new asset class. No one has fit any Garch model on that before, and you want to build a brand new GARCH model. The fitting process is going to be exactly the same as how you build any other ARIMA X model? Right? So last week we talked about how  
The residuals on the 1st stage of ARIMAX aren't supposed to be white noise. The residuals on the second stage are white noise. How do we fit the second stage the same way that you fit any other ARIMA model   
1. You want to ensure that you minimize the AICc, and   
2. You want to ensure that you pass the Ljung box test.   
3. If not, then you just continue increasing the order of the model, so to speak.   
So the same will apply for building a GARCH model

What is the intuition of GARCH in modeling volatility clustering, and mean reversion The existence of volatility clustering and volatility mean reversion in the data for equities normally developed.  
  
Normally developed equity markets GARCH 1,1 possibility of higher GARCH model for more volatile asset classes or emerging markets or frontier markets..

