BITS Pilani

DIGITAL

Excellence made yours

BITS Pilani
DIGITAL
Excellence made yours

Probability and Statistics

Week 1 Live Session

Dr. Pushpendra Gupta

Week 1 Recap

BITS Pilani DIGITAL
Excellence made yours

📊 Central Tendency	→	Mean, Median, Mode
📏 Variability Measures	→	Variance, Standard Deviation, IQR
📈 Data Visualization	→	Histograms, Box Plots, Pie Charts, Scatter Plots
🎯 Distribution Analysis	→	Chebyshev's Inequality, Skewness, Correlation

Today’s Focus

- 🧠 AI Applications of Statistical Concepts
- 🎯 Chebyshev's Inequality in Practice
- ⚖️ Choosing the Right 'Average’
- 📊 Skewness: What is your data is really saying
- 💻 Hands-on Visualization using Google Colab

AI applications of Statistical Concept

BITS Pilani DIGITAL
Excellence made yours

🤖 Your AI Model in Production:

Trained on clean, well-behaved data
Deployed in the messy real world
New data that doesn't follow training patterns

❓ Critical Questions:
How do you detect anomalies?
What if data distribution is unknown?
How do you maintain model reliability?

💡 Answer:
Statistical Foundations!

🎯 Chebyshev's Inequality in Practice

BITS Pilani DIGITAL
Excellence made yours

🔒 The Universal Guarantee:

For ANY dataset, at least $\left(1 - \frac{1}{k^2}\right)$ of data lies within k standard deviations

📊 No assumptions about distribution shape needed!

$k = 2$: At least 75% of data within 2σ
$k = 3$: At least 89% of data within 3σ

🎯 Perfect for AI because:

Real-world data is often non-normal

New data distributions are unknown

Normal Distribution


75%
25%
μ ± 2σ
(μ, σ)

Real-world data

$\left(1-\frac{1}{1.5^2}\right)\times 100$

$= 55.55\%$

$\left(1-\frac{1}{k^2}\right)\times 100\%$ data-ல் ல

K S.D

$\bar{X} = \mu = \text{Mean}, \quad \sigma = (S) = \sqrt{\text{Var}}$

$PDF = f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$

$\mu = 1150, \quad \sigma = 15$

3.14 2.72

$P(a < x < b) = \int_a^b f(x)\,dx$

$a = 100 \quad b = 1300$

$P(X = 150.2345)$

$f(x) = x^2$

$f(3) = 3^2 = 9$

🎯 Chebyshev in Action - Anomaly Detection

BITS Pilani DIGITAL
Excellence made yours

🏦 Fraud Detection System:
Monitor transaction amounts for each user
No assumptions about spending patterns
Use Chebyshev bounds to flag unusual transactions

📊 Implementation:
Calculate mean and σ for user's historical data
Flag transactions beyond 2.5σ as "suspicious"
Guaranteed to catch extreme anomalies

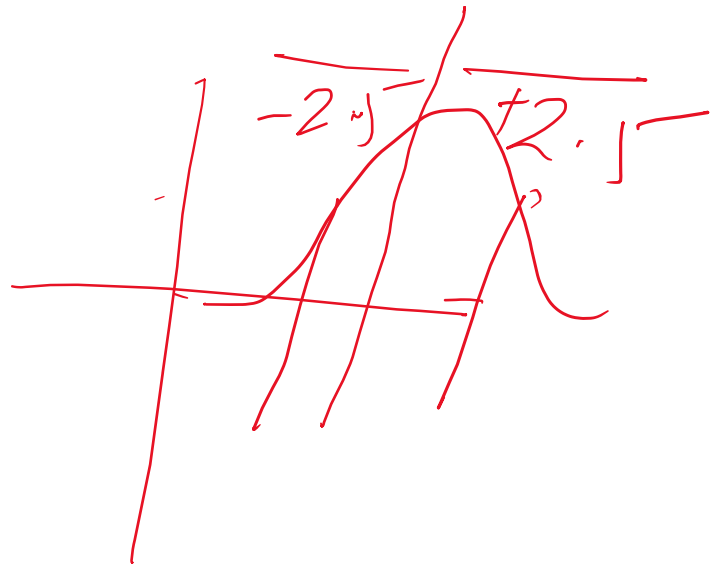
💼 Business Impact:
Reduce false positives (incorrectly flagged as fraudulent)
Catch 84%+ of extreme cases automatically

K^2 $1-\frac{1}{K^2}$ (5a)

$\frac{K^2-1}{K^2}$

50 Engines

2.5 ω



S	Business	✓
₹10	20K	
₹100	30K	
₹10	40K	

Mean μ — 150

α — 50

1.5

2.5

$$\frac{1.75 - 1.2 / }{2.5}$$

$$150 \pm 2.5 \times 50$$

$$= 150 \pm 125.0$$

$$= ₹275, \; 25$$

Thank You