



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



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

 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:

$$16\% \rightarrow$$

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

$\mu - 3\sigma$, $\mu + 3\sigma$

$$\left(1 - \frac{1}{2.5^2}\right) \times 100\%$$



Perfect for AI because:

100%

Real-world data is often non-normal

New data distributions are unknown

16%

= 89%

Normal Distribution



✓
 $\mu \pm 2\sigma$

✓
 (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-4-e w

K S.D

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

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

$$\mu = 1150, \quad \sigma = \underline{150}$$

$$3.14$$

$$2.72$$

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

$$P(x = 150.2345) \quad f(x) = x^2$$

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



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"

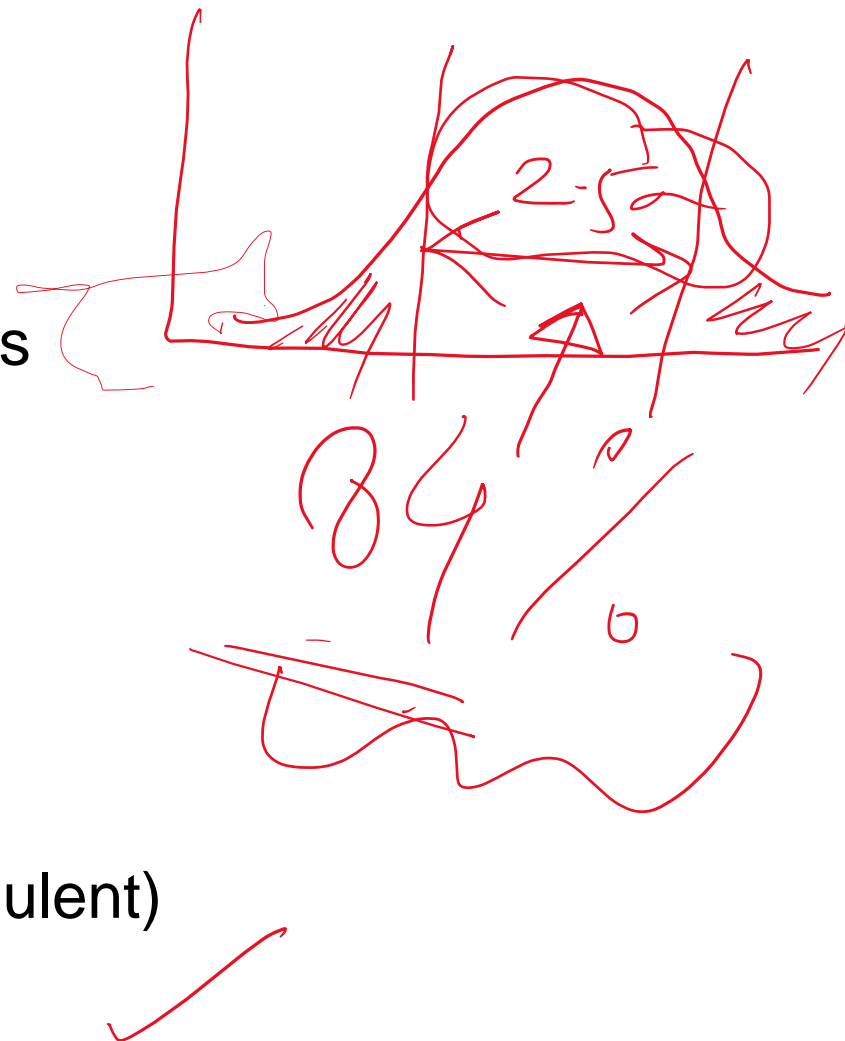
Guaranteed to catch extreme anomalies



Business Impact:

Reduce false positives (incorrectly flagged as fraudulent)

Catch 84%+ of extreme cases automatically



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

$$\left(\frac{K^2 - 1}{K^2} \right)$$



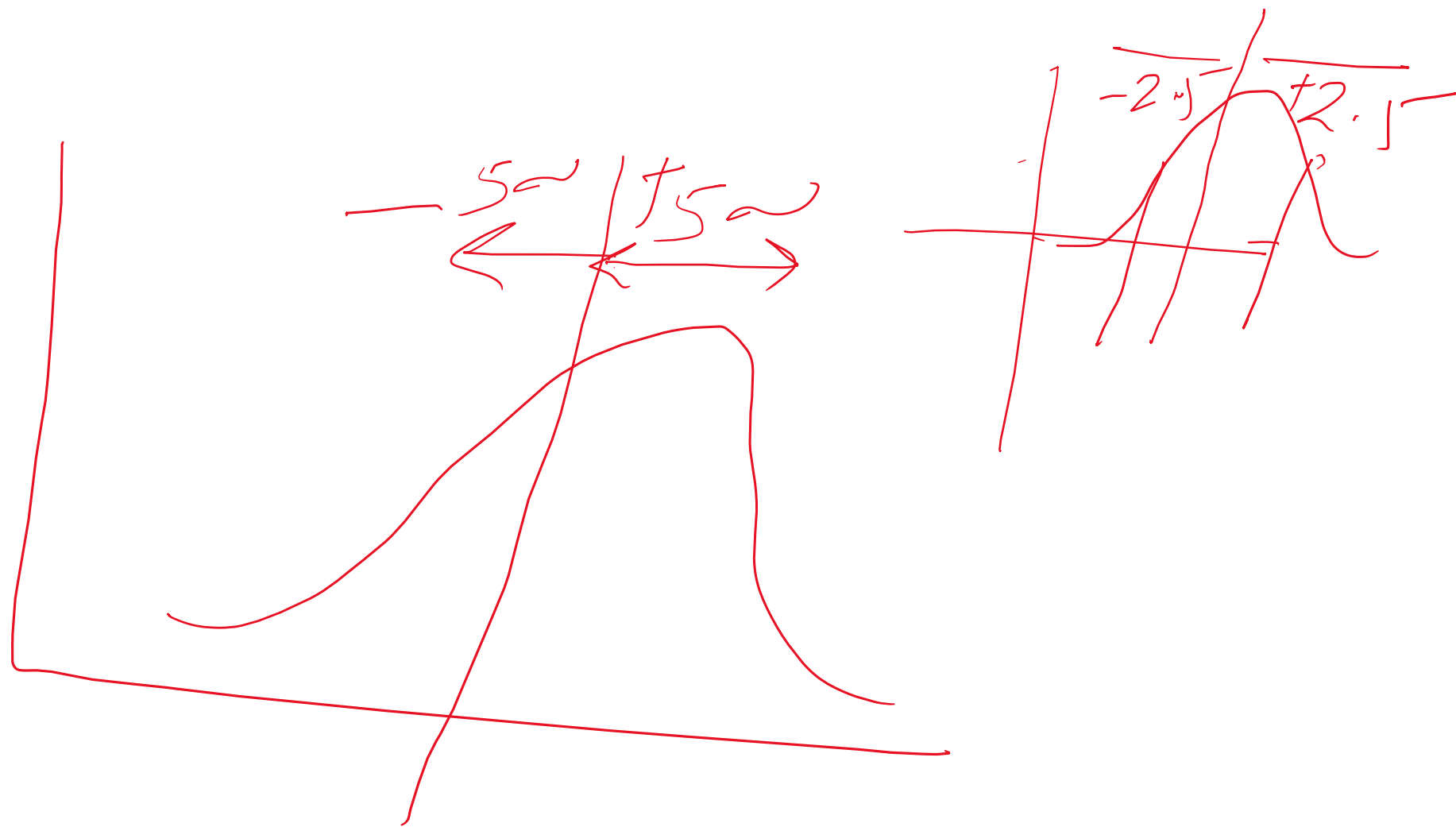
X

100%

51st Cy

50 Engines

2.5a



S	Business	✓
---	----------	---

£10	20K	
-----	-----	--

£100	30K	
------	-----	--

£10	40K	
-----	-----	--

Mean μ = 150
 σ = 50

1.5 σ

2.5 σ

150
 - 125

 25

150 + 2.5 \times 50

= 150 \pm 125.0
 = (275, 25)

Thank You

The content provided in this digital training programme has been developed by professors / external vendors. All intellectual property rights in the content, including but not limited to video, text, slides, graphics, and audio materials, are owned exclusively by BITS Pilani ("BITS"), unless otherwise expressly stated. This programme may also include licensed or contributed works from third-party experts, faculty, and collaborating institutions, all of whom are duly acknowledged where appropriate.

All product names, logos, brands, and entity names referenced within this content are the property of their respective trademark owners. The inclusion of such names, logos, or brands does not imply any affiliation with, or endorsement by, these entities.

This material is intended solely for personal learning and academic purposes. Except for content that is explicitly confirmed to be in the public domain, no part of this content may be reproduced, copied, modified, distributed, republished, uploaded, posted, publicly displayed, or transmitted in any form or by any means – electronic, mechanical, photocopying, recording, or otherwise – without the prior express written consent of BITS Pilani or the applicable rights holder.

Any unauthorized use, commercial exploitation, resale, or public sharing of this content in any manner constitutes a violation of applicable intellectual property laws. BITS Pilani reserves the right to pursue all appropriate legal remedies, including but not limited to civil action and criminal prosecution under the Copyright Act, 1957, and other relevant statutes in force in India.

By accessing or using this digital training programme, you agree to comply with the terms set forth in this disclaimer.