



Table of Contents — Full Conversation Guide

1. Probability Foundations & Discrete Distributions

1.1 Binomial Distribution Expansion

- Writing out the Binomial PMF explicitly
- Interpreting binomial coefficients
- Clarifying notation vs implementation

1.2 Two-Stage Random Models

- Knowledge vs guessing framework
- Conditional distributions
- Constructing joint distributions
- Understanding $Y = N + Z$

1.3 Translating Math to Python

- PMFs as lambda functions
 - \wedge vs $\star\star$ in Python
 - Manual PMF construction vs `scipy.stats`
 - Naming conventions and clarity
-

2. Conditional Probability & Joint Modeling

2.1 Constructing Joint Distributions

- Why probabilities multiply, not add
- $P(N = n, Z = z) = P(N = n)P(Z = z | N = n)$

2.2 Conditioning on Events

- Interpreting $P(N < 10 | Y \geq T)$
- Separating numerator vs denominator
- Common logical errors in conditioning

2.3 Debugging Probability Code

- Accumulation vs overwriting
- Shadowing built-ins (`sum`)

- Normalization sanity checks
-

3. Random Number Generation & Monte Carlo

3.1 Linear Congruential Generators (LCG)

- Definition of LCG
- Role of a, b, m
- Period and reproducibility

3.2 Hull–Dobell Theorem

- Full-period conditions
- Prime factor constraints
- Practical parameter selection
- Worked examples

3.3 Uniform Sampling via Scaling

- Integer PRNG \rightarrow Uniform(0,1)
- Why division by modulus works
- Why the upper bound is never 1

3.4 Monte Carlo Estimation

- Correct Monte Carlo estimators
 - Law of Large Numbers intuition
 - Common estimator mistakes
-

4. Concentration Inequalities & Statistical Guarantees

4.1 What “Exponential Concentration” Means

- Sub-Gaussian vs sub-Exponential
- Bernstein-style bounds
- Two-regime behavior

4.2 Weaker (Polynomial) Concentration

- Chebyshev inequality
- Role of finite variance
- Why finite moments are not enough

4.3 Functional Difficulty Ladder

- Means vs variances vs higher moments
- Tail amplification
- Why squaring hurts concentration

4.4 Exam Strategy & Heuristics

- 10-second decision tree
 - Common traps
 - One-line rules to memorize
-

5. Hoeffding Inequality in Practice (Spam Classification)

5.1 Problem Setup

- SMS spam dataset structure
- Conditional probabilities from data

5.2 Empirical Estimators

- Indicator variables
- Filtering by events
- Correct denominator selection

5.3 Confidence Intervals

- Hoeffding bounds
- Interpreting confidence levels
- Why rare events widen intervals

5.4 Debugging Python Errors

- 'float' object is not callable
 - Shadowed built-ins
 - Kernel restart strategy
-

6. Pandas & Data Handling Essentials

6.1 Counting Observations

- Total rows

- Conditional counts
- Non-missing values

6.2 Unique Values

- `nunique()`
- `value_counts()`
- `unique()`

6.3 Common DataFrame Pitfalls

- Column selection syntax
 - Boolean indexing
 - Efficient filtering
-

7. Markov Chains — Theory & Implementation

7.1 Transition Graph Interpretation

- Reading state diagrams
- Directional probabilities
- Self-loops and transitions

7.2 Transition Matrix Construction

- State ordering
 - Row-stochastic matrices
 - Cross-checking matrices
-

8. Markov Chain Properties

8.1 Irreducibility

- Communication classes
- Reducible vs irreducible chains
- Applying definitions to Chains A–D

8.2 Periodicity & Aperiodicity

- Definition of period
- GCD of return times
- Role of self-loops

- Correct vs incorrect reasoning
- Chain-by-chain validation

8.3 Stationary Distributions

- What “finite” means
 - Existence vs uniqueness
 - When reducible chains still have stationary distributions
-

9. Computing Stationary Distributions

9.1 Power Iteration Method

- Why it works
- When it fails
- Convergence assumptions

9.2 Linear System Method

- Solving $\pi P = \pi$
- Constraint handling
- Numerical stability

9.3 Comparing Methods

- Strengths and weaknesses
 - When to prefer each
-

10. Reversibility of Markov Chains

10.1 Definition

- Detailed balance condition
- Intuition via probability flow

10.2 Testing Reversibility

- Algorithmic check
- Numerical tolerance
- Interpretation of failures

10.3 Applying to Chains A–D

- Correct implementation
 - Logical interpretation of results
-

11. Frequency Dictionaries & MLE for Markov Chains

11.1 Frequency Mapping

- Purpose of `makeFreqDict`
- Correct input structure

11.2 Extracting Transitions from DataFrames

- Why `df['from', 'to']` fails
- Correct `zip(df['from'], df['to'])` usage

11.3 Maximum Likelihood Estimation

- Counting transitions
 - Row normalization
 - Avoiding NaNs
-

12. Debugging & Python Best Practices (Meta Section)

12.1 Common Python Gotchas

- Shadowing built-ins
- Callable vs non-callable errors

12.2 Numerical Stability

- Floating-point cleanup
- Normalization safeguards

12.3 Exam-Safe Coding Habits

- Naming conventions
 - Sanity checks
 - Minimal assumptions
-

13. Conceptual Meta-Lessons

13.1 What the Assignments Are Really Testing

- Probability reasoning over computation
- Structural understanding

13.2 How to Think Under Exam Pressure

- Pattern recognition
 - Eliminating distractors
 - Interpreting wording precisely
-

14. Summary Cheat Sheets

- Concentration inequalities
- Markov chain properties
- Monte Carlo estimation
- PRNG design rules