Lecture 9

5. Summation Tricks:

a. For expectation: \( E(X) = \sum_x x p_X(x) \)

i. For probabilities involving factorial of random variable value

- Incorporate random variable value into the factorial.
  - If the factorial is in the numerator, move argument up by one: Negative binomial
  - If the factorial is in the denominator, move argument down by one: Binomial, Poisson

- Re-parameterize the remainder of the quantities
- Pull out quantities not changing with random variable value
- Identify sum recognizable as summing to 1

ii. For probabilities involving powers of the model parameters:

- Geometric
  - Note that integer times this power represents a derivative
  - Pull out model parameter quantities not involved in the derivative.
  - Pull derivative outside the sum
  - Recognize the sum as a differentiable function
  - Perform the derivative.
b. Second Moment:
   i. Sometimes is easier to calculate expectation of $X(X - 1)$ if factorial is in denominator.
   ii. Or $X(X + 1)$ if factorial is in numerator.

c. Moment and Probability Generating Function $E(\exp(tX))$ and $E(t^X)$
   i. These involve the expectation of a quantity raised to power $X$
      • $t$ for probability generating function
      • $\exp(t)$ for mgf
   ii. Most of the probability functions under consideration involve a quantity raised to the power $x$.
   iii. Multiply these to get a single quantity raised to the power $x$.
   iv. Recognize this as the probability function for a distribution of the same form with a different parameter.

WMS: 3.11

6. Probability Inequalities
   a. Markov Inequality: If $P(X \geq 0) = 1$, then for all $t > 0$,
      $P(X \geq t) \leq E(X)/t$.
      i. Марков (Markoff)
ii. Proof: Split sum according to whether $x$ is at least as great as $t$.
\[ E(X) = \sum x \, x p_X(x) = \sum_{x < t} x p_X(x) + \sum_{x \geq t} x p_X(x) \]
- Part with lower values is positive, by positivity; throw it away:
  \[ E(X) \geq \sum_{x \geq t} x p_X(x) \].
- Bound sum below by lower bound on $x$:
  \[ E(X) \geq \sum_{x \geq t} t \, p_X(x) \].
- Factor out $t$:
  \[ E(X)/t \geq \sum_{x \geq t} t \, p_X(x) \].

b. Tchebysheff Inequality: \[ P \left( (X - E(X))^2 \geq t \right) \leq V(X)/t \],
as can be seen by applying Markov’s inequality to \((X - E(X))^2\).

i. Чебышёв (Chebyshev, Chebychov, Chebyshov; or Tchebychev, Tchebycheff (French transcriptions); or Tschebyschev, Tschebyschef, Tschebyscheff (German transcriptions); Čebyčev. (from Wikipedia).

c. We can use Tchebysheff inequality to show \( V(X) = 0 \) if and only if there exists \( c \) such that \( P(X = c) = 1 \).

i. Suppose \( P(X = c) = 1 \).
- Then \( E(X) = c \)
- Also \( V(X) = \sum_{x=c} (x - c)^2 p_X(x) = 0 \).

ii. Suppose \( V(X) = 0 \).
• By Tchebysheff inequality, \( \Pr(|X - \mathbb{E}(X)| \geq t) \leq \frac{\text{V}(X)}{t^2} \) for all \( t > 0 \).

• Let \( A = \{X \neq \mathbb{E}(X)\} = \{|X - \mathbb{E}(X)| > 0\} \)

• Let \( A_n = \{|X - \mathbb{E}(X)| \geq 1/n\} \).

• Then \( A = \bigcup A_n \)

• Then \( \Pr(A) \leq \sum_{n=1}^{\infty} \Pr(A_n) = 0 \).

WMS: 4.1

IV. Continuous Distributions

A. Introduction to Continuous Distributions

1. Definition and Examples

   a. Recall distribution function \( F_X(x) = \Pr(X \leq x) \).

   b. A continuous distn is one in which the distribution function is given an integral \( F_X(x) = \int_{-\infty}^{x} f_X(y) \, dy \).

   i. To calculate the probability associated with a particular range of values, integrate a function called a probability density function over the range.

   c. Such a variable could take on any value in a range of real numbers.

   d. Examples
1. Change in blood pressure or weight
   ii. Times until an event happens
   iii. Asset price changes, asset yields.

2. Are any random variables really continuous?
   a. Physical measurements are
      i. limited in possible values by precision of measuring apparatus
      ii. limited by quantum nature of physical reality
   b. Financial quantities are limited by discreteness in measurement units.
      i. Ex. US markets used to measure stock prices in multiples of $0.125.
   c. If distribution is truly discrete, with sufficiently fine inter-point distance, continuous distributions will be an adequate approximation.

3. Information about probabilities is contained in the distribution function.
   a. As before, $F_X(x) = P(X \leq x)$.
   b. Shares monotonicity, limits at $\pm\infty$ with discrete case.
      i. $F_X(x)$ is non-decreasing
ii. $\lim_{x \to \infty} F_X(x) = 1$

iii. $\lim_{x \to -\infty} F_X(x) = 0$.

c. Consequence: $\int_{-\infty}^{\infty} f_X(x) \, dx = 1$

i. Can restrict to any range outside of which $f_X(x) = 0$

WMS: 4.2

4. The density is the derivative of the distribution function.

a. $dF_X(x_1)/dx_1 = f_X(x_1)$, if $f_X$ is continuous at $x_1$,

i. Take:

• $x_2 > x_1$,

• and if $x_2$ is close to $x_1$,

ii. Then

• $F_X(x_2) - F_X(x_1) = P(x_1 < X \leq x_2) = \int_{x_1}^{x_2} f_X(y) \, dy \approx (x_2 - x_1)f_X(x_1)$

• Hence $(F_X(x_2) - F_X(x_1))/(x_2 - x_1) \approx f_X(x_1)$

• Hence $dF_X(x_1)/dx_1 = f_X(x_1)$: Fundamental theorem of calculus.

b. Uniform Example:

i. $f_X(x) = 1$ for $x \in (0, 1)$, and equal 0 elsewhere.

• Called “Uniform on $(0, 1)$”.


ii. Then

\[ F_X(x) = \begin{cases} 
0 & \text{if } x \leq 0 \\
 x & \text{if } x \in (0, 1) \\
1 & \text{if } x \geq 1 
\end{cases} \]

iii. Then

\[ F_X'(x) = \begin{cases} 
0 & \text{if } x < 0 \\
1 & \text{if } 1 \in (0, 1) \\
0 & \text{if } x > 1 \\
\text{undefined} & \text{if } x \in \{0, 1\} 
\end{cases} \]