

<p>RANDOM VARIABLE DEFINITION</p> <p><i>What is a random variable?</i></p>	<p>DISCRETE VS CONTINUOUS</p> <p><i>What's the difference between discrete and continuous random variables?</i></p>
<p>PROBABILITY MASS FUNCTION</p> <p><i>What is a probability mass function (pmf)?</i></p>	<p>PROBABILITY DENSITY FUNCTION</p> <p><i>What is a probability density function (pdf)?</i></p>
<p>CUMULATIVE DISTRIBUTION FUNCTION</p> <p><i>What is a cumulative distribution function (cdf)?</i></p>	<p>CDF PROPERTIES</p> <p><i>What are the key properties of any cumulative distribution function?</i></p>
<p>CONDITIONAL PROBABILITY</p> <p><i>What does conditional probability tell us?</i></p>	<p>INDEPENDENCE</p> <p><i>When are two events independent?</i></p>
<p>IDENTICALLY DISTRIBUTED</p> <p><i>What does it mean for random variables to be identically distributed?</i></p>	<p>TRANSFORMATION OF VARIABLES</p> <p><i>How do we find the distribution of a function of a random variable?</i></p>

<p>A discrete random variable can only take on countably many values (like integers), often resulting from counting. A continuous random variable can take on any value in an interval, often resulting from measurement. The key difference is whether the variable has "gaps" between possible values.</p>	<p>A random variable is a function that assigns a real number to each outcome in a sample space. It's essentially a rule that converts the results of a random experiment into numerical values that we can work with mathematically.</p>
<p>A pdf describes the probability distribution of a continuous random variable. Unlike a pmf, the pdf value at a point is not a probability - instead, probabilities are found by integrating the pdf over intervals. The pdf must be non-negative and integrate to 1.</p>	<p>A pmf describes the probability distribution of a discrete random variable. It gives the probability that the variable equals each specific value. The pmf must be non-negative and all probabilities must sum to 1.</p>
<p>Every cdf must satisfy: (1) values between 0 and 1, (2) non-decreasing, (3) right-continuous, (4) approaches 0 as <math>x</math> approaches negative infinity, and (5) approaches 1 as <math>x</math> approaches positive infinity.</p>	<p>The cdf gives the probability that a random variable is less than or equal to a given value. It works for both discrete and continuous variables, always increases from 0 to 1, and is right-continuous. It completely characterizes the distribution.</p>
<p>Two events are independent if knowing that one occurred doesn't change the probability of the other occurring. Mathematically, A and B are independent if <math>P(A B) = P(A)</math>, or equivalently, <math>P(AB) = P(A)P(B)</math>.</p>	<p>Conditional probability <math>P(A B)</math> gives the probability of event A occurring given that we know event B has occurred. It updates our knowledge based on new information and is calculated as <math>P(AB)/P(B)</math> when <math>P(B) &gt; 0</math>.</p>
<p>For <math>Y = g(X)</math>, we can find the distribution of Y using the transformation method. For continuous variables, we often use the inverse transformation and the Jacobian. For discrete variables, we find where each y-value comes from in the original distribution.</p>	<p>Two random variables are identically distributed if they have the same probability distribution - meaning they have the same cdf (or pmf/pdf). Note that identically distributed variables are not necessarily equal; they just follow the same probability laws.</p>

<p>EXPECTED VALUE</p> <p><i>What is the expected value of a random variable?</i></p>	<p>VARIANCE</p> <p><i>What does variance measure?</i></p>
<p>STANDARD DEVIATION</p> <p><i>What is standard deviation and why is it useful?</i></p>	<p>MOMENTS</p> <p><i>What are moments of a distribution?</i></p>
<p>MOMENT GENERATING FUNCTION</p> <p><i>What is a moment generating function and why is it useful?</i></p>	<p>MGF PROPERTIES</p> <p><i>What are the key properties of moment generating functions?</i></p>
<p>LINEAR COMBINATIONS</p> <p><i>How do expectations behave with linear combinations?</i></p>	<p>VARIANCE OF LINEAR COMBINATIONS</p> <p><i>How does variance behave with linear combinations?</i></p>
<p>INDEPENDENCE VS CORRELATION</p> <p><i>What's the relationship between independence and correlation?</i></p>	<p>LAW OF TOTAL PROBABILITY</p> <p><i>What is the law of total probability?</i></p>

<p>Variance measures how spread out a distribution is around its mean. It's the expected value of the squared deviations from the mean: <math>\text{Var}(X) = E[(X - \mu)^2]</math>. Larger variance means more spread; variance of zero means the variable is constant.</p>	<p>The expected value (or mean) is the long-run average value of a random variable. For discrete variables, it's the sum of each value times its probability. For continuous variables, it's the integral of <math>x</math> times the pdf. It represents the "center" of the distribution.</p>
<p>Moments are expected values of powers of a random variable. The first moment is the mean, the second central moment is the variance. Higher moments describe the shape of the distribution - the third moment relates to skewness, the fourth to kurtosis.</p>	<p>Standard deviation is the square root of variance. It measures spread in the same units as the original variable, making it easier to interpret than variance. A small standard deviation means values cluster near the mean; a large one means they're more dispersed.</p>
<p>If the mgf exists: (1) <math>M(0) = 1</math> always, (2) the <math>n</math>th derivative at <math>t=0</math> gives the <math>n</math>th moment, (3) <math>M(t)</math> uniquely determines the distribution, (4) for <math>Y = aX + b</math>, <math>M_Y(t) = e^{bt}M_X(at)</math>, and (5) convergence of mgf implies convergence of distributions.</p>	<p>The moment generating function (mgf) is <math>M(t) = E[e^{tX}]</math>. If it exists in a neighborhood of zero, it uniquely determines the distribution. It's also useful for proving convergence results.</p>
<p>For <math>\text{Var}(aX + b)</math>, we get <math>a^2\text{Var}(X)</math> - the constant <math>b</math> disappears and <math>a</math> gets squared. For sums, <math>\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\text{Cov}(X, Y)</math>. If <math>X</math> and <math>Y</math> are independent, the covariance term is zero.</p>	<p>Expectation is linear: <math>E[aX + bY + c] = aE[X] + bE[Y] + c</math> for any constants <math>a, b, c</math>. This holds regardless of whether <math>X</math> and <math>Y</math> are independent. However, variance is not linear unless variables are independent.</p>
<p>If events <math>B_1, B_2, \dots, B_n</math> partition the sample space, then <math>P(A) = \sum P(A B_i)P(B_i)</math>. This lets us calculate the probability of <math>A</math> by conditioning on different scenarios and weighing each by how likely that scenario is.</p>	<p>Independence implies zero correlation (uncorrelated), but zero correlation doesn't imply independence. Independence is a stronger condition - it means the variables don't depend on each other in any way, while zero correlation only means they're not linearly related.</p>

<p>BAYES' THEOREM</p> <p><i>What is Bayes' theorem and when is it used?</i></p>	<p>PROBABILITY VS DENSITY</p> <p><i>What's the key difference between probability and density for continuous variables?</i></p>
<p>TRANSFORMATION TECHNIQUE</p> <p><i>What's the general approach for transforming continuous random variables?</i></p>	<p>COMMON EXPECTATION FORMULAS</p> <p><i>What are some useful expectation formulas to remember?</i></p>
<p>EXISTENCE OF MOMENTS</p> <p><i>When do moments exist?</i></p>	<p>DISTRIBUTION UNIQUENESS</p> <p><i>When does a moment generating function uniquely determine a distribution?</i></p>
<p>PROBABILITY AXIOMS</p> <p><i>What are the three fundamental axioms of probability?</i></p>	<p>SAMPLE SPACE</p> <p><i>What is a sample space?</i></p>
<p>FUNCTION OF RANDOM VARIABLE</p> <p><i>What makes a function of a random variable also a random variable?</i></p>	<p>CONTINUOUS PROBABILITY PROPERTIES</p> <p><i>What special properties do continuous random variables have?</i></p>

<p>For continuous variables, the pdf value at a point is not a probability - individual points have probability zero. Instead, probabilities are areas under the curve (integrals). The pdf tells us about the relative likelihood or "concentration" of values in different regions.</p>	<p>Bayes' theorem gives <math>P(B A) = \frac{P(A B)P(B)}{P(A)}</math>. It's used to "reverse" conditional probabilities - if we know <math>P(A B)</math> but want <math>P(B A)</math>. It's fundamental to updating beliefs based on new evidence.</p>
<p><math>E[aX + b] = aE[X] + b</math>, <math>E[X^2] = \text{Var}(X) + [E(X)]^2</math>, <math>E[XY] = E[X]E[Y]</math> if independent, and for finding <math>E[g(X)]</math> you can either use the definition with the original pdf or transform to find the pdf of <math>Y = g(X)</math> first.</p>	<p>For <math>Y = g(X)</math> where <math>g</math> is monotonic: (1) find the inverse function <math>X = g^{-1}(Y)</math>, (2) differentiate to get <math>dx/dy</math>, (3) substitute into <math>f_Y(y) = f_X(g(y)) dx/dy </math>. For non-monotonic functions, split into monotonic pieces and add the contributions.</p>
<p>If the mgf exists in some open interval containing zero, then it uniquely determines the distribution. Even if all moments exist, they might not uniquely determine the distribution unless additional conditions are met (like bounded support or Carleman's condition).</p>	<p>The <math>n</math>th moment <math>E[X^n]</math> exists if the integral (or sum) converges. Heavy-tailed distributions may not have all moments. If the <math>n</math>th moment exists, all lower moments also exist. The mgf existing implies all moments exist, but having all moments doesn't guarantee the mgf exists.</p>
<p>The sample space <math>S</math> is the set of all possible outcomes of a random experiment. It must be exhaustive (include all possibilities) and mutually exclusive (outcomes don't overlap). Events are subsets of the sample space.</p>	<p>Kolmogorov's axioms: (1) <math>P(A) \geq 0</math> for any event <math>A</math>, (2) <math>P(S) = 1</math> where <math>S</math> is the sample space, and (3) for disjoint events <math>A, A', \dots</math>, <math>P(A \cup A' \cup \dots) = P(A) + P(A') + \dots</math>. All other probability rules follow from these axioms.</p>
<p>For continuous random variables: (1) <math>P(X = \text{any specific value}) = 0</math>, (2) <math>P(a \leq X \leq b) = P(a &lt; X &lt; b) = P(a \leq X &lt; b) = P(a &lt; X \leq b)</math>, (3) the cdf is continuous, and (4) probabilities are found by integrating the pdf.</p>	<p>If <math>X</math> is a random variable and <math>g</math> is a function, then <math>Y = g(X)</math> is also a random variable. The key is that <math>Y</math> assigns a number to each outcome in the sample space through the composition: <math>\text{outcome} \rightarrow X(\text{outcome}) \rightarrow g(X(\text{outcome})) = Y(\text{outcome})</math>.</p>