Problem solving strategies Part 1

Tara Templin

Approaches to “prove/show” stuff with math

Direct proof (if a, then b)

Related: **Proof by cases** (enumerate all the cases and show true everywhere)

Proof by induction (show true for n=1, if true for n=k-1, then also n=k)

Proof by contradiction (suppose if not a, then b. Show ridiculous. Thus if a, then b)

Proof by contrapositive (if not b, then not a)

Proof by construction (e.g. counter example. E[X] always defined? No, cauchy)

Example Problem: Direct Proof

Example: Show that the sum of any two odd integers is even

Example Problem: Direct Proof

Example: Show that the sum of any two odd integers is even

Let a and b be any two odd integers, so $a = 2k + 1$ and $b = 2l + 1$ for any k,l integers.

$A + b = (2k + 1) + (2l + 1) = 2k + 2l + 2 = 2 * (k + l + 1)$

Example Problem: Proof by Contradiction

Prove that if a^2 is even, then a is even, for all integers a

Example Problem: Proof by Contradiction

Prove that if a^2 is even, then a is even, for all integers a

Suppose by contradiction that the proposition is not true.

That is, there is a number a such that a^2 is even but a is not even -> a is odd.

So $a = 2k + 1$ for some integer k

This implies that $a^2 = (2k + 1)^2 = 4k^2 + 4k + 1 = 2(2k^2+2k) +1$

That is, a^2 is odd. This is a contradiction to the assumption that a^2 is even.

So the supposition must be wrong. The proposition is true.

Approach for Proof by Induction

To prove something like, P(n) is true for n = 0, 1, 2, 3, … (integers) it is enough to prove:

1) The base case: prove that P(0) is true
2) The inductive case: if P(n-1) is true (the “inductive hypothesis”), then prove that P(n) must also be true.

Example Problem: Proof by Induction

Show that $\Pr(A_1 \ldots A_n) = \Pr(A_1)\Pr(A_2|A_1)\Pr(A_3|A_1A_2)\ldots\Pr(A_n|A_1\ldots A_{n-1})$

Example Problem: Proof by Induction

Show that $Pr(A_1 ... A_n) = Pr(A_1)Pr(A_2|A_1)Pr(A_3|A_1A_2)...Pr(A_n|A_1...A_{n-1})$

Base case: $Pr(A_1A_2) = Pr(A_1)Pr(A_2|A_1)$ [from the definition of conditional prob]

Now set up the Inductive hypothesis: Assume the statement is correct for $k = n - 1$

Example Problem: Proof by Induction

Show that $Pr(A_1...A_n) = Pr(A_1)Pr(A_2|A_1)Pr(A_3|A_1A_2)...Pr(A_n|A_1...A_{n-1})$

Base case: $Pr(A_1A_2) = Pr(A_1)Pr(A_2|A_1)$ [from the definition of conditional prob]

Inductive hypothesis: Assume the statement is correct for $k = n - 1$

$Pr(A_1...A_{n-1}) = Pr(A_1)Pr(A_2|A_1)...Pr(A_{n-1}|A_1...A_{n-2})$

Let "A_1" = $A_1...A_{n-1}$ and "A_2"=A_n

$Pr(A_1...A_n) = Pr(A_1...A_{n-1})\ Pr(A_n|A_1...A_{n-1})$ [definition of conditional prob]

$Pr(A_1...A_n) = Pr(A_1)Pr(A_2|A_1)Pr(A_3|A_1A_2)...Pr(A_n|A_1...A_{n-1})$ [inductive hypothesis]

Common Summations to Know

Very common to work with sums in discrete statistics

Recognize ways to simply or manipulate the summation when you can

$$\sum_{n=s}^{t} C \cdot f(n) = C \cdot \sum_{n=s}^{t} f(n)$$ (distributivity)

$$\sum_{n=s}^{t} f(n) \pm \sum_{n=s}^{t} g(n) = \sum_{n=s}^{t} (f(n) \pm g(n))$$ (commutativity and associativity)

$$\sum_{n=s}^{t} f(n) = \sum_{n=s+p}^{t+p} f(n-p)$$ (index shift)

More on index shifts in a moment!

Source: https://en.wikipedia.org/wiki/Summation

Common Summations to Know

$$\sum_{n=s}^{t} f(n) = \sum_{n=s}^{j} f(n) + \sum_{n=j+1}^{t} f(n)$$ (splitting a sum, using associativity)

$$\sum_{n=a}^{b} f(n) = \sum_{n=0}^{b} f(n) - \sum_{n=0}^{a-1} f(n)$$ (a variant of the preceding formula)

$$\sum_{i=k_0}^{k_1} \sum_{j=l_0}^{l_1} a_{i,j} = \sum_{j=l_0}^{l_1} \sum_{i=k_0}^{k_1} a_{i,j}$$ (commutativity and associativity, again)

$$\sum_{k \le j \le i \le n} a_{i,j} = \sum_{i=k}^{n} \sum_{j=k}^{i} a_{i,j} = \sum_{j=k}^{n} \sum_{i=j}^{n} a_{i,j} = \sum_{j=0}^{n-k} \sum_{i=k}^{n-j} a_{i+j,i}$$ (another application of commutativity and associativity)

$$\sum_{n=0}^{2t+1} f(n) = \sum_{n=0}^{t} f(2n) + \sum_{n=0}^{t} f(2n+1)$$ (splitting a sum into its odd and even parts, and changing the indices)

$$\left(\sum_{i=0}^{n} a_i\right)\left(\sum_{j=0}^{n} b_j\right) = \sum_{i=0}^{n} \sum_{j=0}^{n} a_i b_j$$ (distributivity)

Source: https://en.wikipedia.org/wiki/Summation

Common Summations to Know

Powers and logarithm of arithmetic progressions [edit]

$$\sum_{i=1}^{n} c = nc$$ for every c that does not depend on i

$$\sum_{i=0}^{n} i = \sum_{i=1}^{n} i = \frac{n(n+1)}{2}$$ (Sum of the simplest arithmetic progression, consisting of the n first natural numbers.)

$$\sum_{i=1}^{n} (2i-1) = n^2$$ (Sum of first odd natural numbers)

$$\sum_{i=0}^{n} 2i = n(n+1)$$ (Sum of first even natural numbers)

There are many series and infinite sums. Let's revisit index shifts with the commonly used "geometric series"

Source: https://en.wikipedia.org/wiki/Summation

Application: Geometric Series

Both of these are infinite geometric series: $\sum_{n=0}^{\infty} ar^n \qquad \sum_{n=1}^{\infty} ar^{n-1}$

The only difference is an index shift.

Remember the solution of a finite/infinite geometric series?

Application: Geometric Series

Both of these are infinite geometric series: $\sum_{n=0}^{\infty} ar^n$ $\sum_{n=1}^{\infty} ar^{n-1}$

The only difference is an index shift.

Remember the solution of a finite/infinite geometric series?

$$\sum_{k=0}^{n-1} ar^k = a\left(\frac{1-r^n}{1-r}\right)$$

$$\sum_{k=0}^{\infty} ar^k = \frac{a}{1-r}, \text{ for } |r| < 1$$

When working with a sum, look to use change of variables to solve, e.g.:

$$\sum_{y=1}^{x}(1-\theta)^{y-1}\theta = \sum_{n=0}^{x-1}(1-\theta)^n\theta = \theta\sum_{n=0}^{x-1}(1-\theta)^n = 1-(1-\theta)^x$$

Approaches to Calculate Expected Values

A common scenario is finding the expected value of an unfamiliar distribution:

- Use basic properties of expected values
  - E.g. Linearity ($E[aX+Y]=aE[X]+E[Y]$)
- Use change of variables to match common distributions with known means
- If the distribution can be transformed into a known discrete PDF, use the fact that discrete PDFs sum to 1
- Remember, $E(f(X))$ does not usually equal $f(E(X))$
  - Use Law of the unconscious statistician
- Define an indicator variable to denote if an event occurs: $E[1(A)] = P(A)$
- Look for independence: $E[XY]=E[X]E[Y]$

Example Problem: Expected Values

Let X be a discrete RV that takes on only nonnegative integer values. Show that $E(X) = \sum_{i=0}^{\infty} Pr(X > i)$ [Rice, 3rd]

From the definition of a discrete probability distribution: $E[X] = \sum x_i * Pr(X=x_i)$

$E[X] = 0 * Pr(X=0) + 1 * Pr(X=1) + 2 * Pr(X=2) + 3 * Pr(X=3) + \dots$

$E[X] = Pr(X=1) + [Pr(X=2) + Pr(X=2)] + [Pr(X=3) + Pr(X=3) + Pr(X=3)] + \dots$

$E[X] = [Pr(X=1)+Pr(X=2) +Pr(X=3)...] + [Pr(X=2)+Pr(X=3)...] + [Pr(X=3)+...] + \dots$

$E[X] = Pr(X >=1) + Pr(X >= 2) + Pr(X >= 3) + \dots = \sum_{i=1}^{\infty} Pr(X >= i) = \sum_{i=0}^{\infty} Pr(X > i)$

Example Problem: Expected Values

Apply the previous result to find the mean of a geometric RV. [Rice, 3rd]

Geometric RV = "The probability distribution of the number $Y = X - 1$ of failures before the first success, supported on the set { 0, 1, 2, 3, ... }"

$\Pr(X > i) = \sum^{\infty}_{j=i+1}[p^*(1-p)^{j-1}]$

$= p^*(1-p)^i \sum^{\infty}_{k=0}[(1-p)^k]$

$= p^*(1-p)^{i} * 1/(1-(1-p)) = (1-p)^i$

Apply result: $E(X) = \sum^{\infty}_{i=0} \Pr(X > i) = \sum^{\infty}_{i=0}(1-p)^i = 1/(1-(1-p)) = 1/p$ [geom series]

Example Problem: Expected Values/Index shifting

We can also use index shifting:

$$\sum_{j=1}^{\infty} P(X \geq j) = \sum_{j=1}^{\infty} \sum_{k=j}^{\infty} P(X = k) = \sum_{k=1}^{\infty} \sum_{j=1}^{k} P(X = k) = \sum_{k=1}^{\infty} kP(X = k) = E[X]$$

First step: probability of a discrete var

Second step: Interchange order of summation bc can sum up however we want.

Third step: Recognizes it’s the definition of an expectation.

Example Problem 2

If $U_1, ..., U_n$ are iid Unif(0, 1) , find $E(U_{(n)} - U_{(1)})$, where $U_{(n)} = \max\{U_1, ..., U_n\}$ and $U_{(1)} = \min\{U_1, ... , U_n\}$. [Rice, 3rd]

Want the expectation of a function of RVs. Two approaches: (1) Could define a new RV, $Z = U_{(n)} - U_{(1)}$, and then use the formula for an expectation and hope it works out.

Example Problem 2

If U_1, ..., U_n are iid Unif(0, 1) , find $E(U_{(n)}−U_{(1)})$, where $U_{(n)} = \max\{U_1, ..., U_n\}$ and $U_{(1)} = \min\{U_1, ... , U_n\}$. [Rice, 3rd]

Want the expectation of a function of RVs. Two approaches: (1) Could define a new RV, $Z = U_{(n)} − U_{(1)}$, and then use the formula for an expectation and hope it works out.

(2) We know expectation is a linear operator, so $E(U_{(n)} − U_{(1)})=E(U_{(n)}) − E(U_{(1)})$. We know a general formula for $f(U_{(n)})$, and $f(U_{(1)})$ but it's still terrible, so let's keep using info. We have U_1, ..., U_n are iid uniform U(0, 1), and that is key.

We know $U_{(k)}$~Beta(k,n+1-k). If X ~ Beta(a,b), then E[X] = a/(a+b)

$E(U_{(n)} − U_{(1)})=E(U_{(n)}) − E(U_{(1)})$ = n/(1+n) - 1/(n+1) = (n-1)/(n+1)

Application: Sums of Random Variables

The sum of exponential random variables is often used in stochastic processes, particularly anything involving queueing.

Imagine a network of computers sending jobs to a compute node or towns sending people to a central hospital.

Each computer or town has a rate parameter of sending compute jobs or people, and the "load" of the compute node or hospital is the sum of those random variables.

The distribution of that sum can be used to estimate wait times for patients.

Application: Order Statistics

One way that order statistics can be used is for modeling government procurement.

There's often a bidding procedure that happens where companies bid for gov't health sector projects and the lowest-cost bid will win.

If the bids are independent of each other (which they often aren't, so you see how covariance can get interesting!), then, for each bid X_i, the distribution of $U = \min\{X_1, . . . , X_n\}$ tells the gov't what distribution of costs to expect for their projects that have project bidding events.