Introduction to Probability & Statistics
Prof. Abhay Gopal Bhatt
Department of Statistics
Indian Statistical Institute Delhi
Week - 5
Lecture - 17
Negative Binomial Distribution

Ki aaj hum Negative Binomial Distribution ya Negative Binomial Random Variables ki taraf dhyan denge; jaise ki Binomial Distribution me tha, humne dekha tha ki ek trial me ya to success hota hai ya failure hota hai, aur hum binaomial (Bernoulli) trial ko baar-baar repeat karte hain, aur uske baad number of trials fix hota tha n, aur un n trials me hum dekhte the ki kitni baar success aaya; Negative Binomial me thoda alag hai lekin core concept wahi hai; Negative Binomial experiment me hum repeated trials karte hain, har trial me parinaam ya to S (success) ya F (failure) hota hai, har trial me success ki probability same hoti hai jise hum p se darshate hain, jiska maan 0 aur 1 ke beech hota hai; ab difference ye hai ki yahan number of trials fix nahi hota hum trials tab tak karte rahenge jab tak hume r successes na mil jaaye; yeh experiment humne pehle r=1 ke special case me dekha hai: ek coin ko baar-baar uchhaalte rahna jab tak pehla head na mil jaaye; yaani trials independent hote hain aur r-th success aane par experiment ruk jata hai; aur hum random variable X define karte hain as the number of failures that precede the r-th success yaani X = r-th success se pehle aaye failures ki sankhya; yahi random variable Negative Binomial Random Variable kehlata hai with parameters r and p, jahan r ek positive integer hota hai aur p (probability of success) 0 aur 1 ke beech; notation: X ~ negative binomial (r, p); ab iski PMF nikaalte hain: X = number of failures before r-th success; X shunya bhi ho sakta hai agar pehle hi r trials me rs success aaye to X=0; X=1,2,3,… koi bhi non-negative integer ho sakta hai; ab P(X=x) ka matlab hai ki r-th success aane se pehle exactly x failures aaye; yaani total number of trials hua x + r; in x+r trials me last trial me zaroor success hota hai kyunki wahi r-th success hai; isliye last slot me S fixed hai; baaki first x+r–1 slots me hume chahiye X failures aur R–1 successes; in slots me R–1 successes aur X failures ko arrange karne ke tareeqe hote hain (x+r–1 choose x) ya equivalently (x+r–1 choose r–1); har aise arrangement me R successes honge probability p^r ke saath, aur X failures honge probability (1–p)^x ke saath; to PMF hota hai: P(X=x) = (x+r–1 choose x) · p^r · (1–p)^x, jahan x=0,1,2,3; isi prakar humne negative binomial distribution ki PMF derive kar li hai.x plus r minus 1 choose r minus 1, p to the power r, 1 minus p to the power x, jabki x values lenge 0,1,2, aur agar x inke alawa koi value leta hai to P(x)=0 hoga; in sab probabilities ka sum agar karein to iska yog bhi 1 aata hai, aur jaise binomial distribution me summation ke liye binomial expansion use hota hai, waise hi yahan summation me negative exponent aata hai isi wajah se ise negative binomial distribution kaha jata hai; expected value aur variance bhi bilkul binomial jaisa hi derive kiya ja sakta hai main direct result likhta hoon: agar X negative binomial (r,p) hai, to expected value

$$E(X) = \frac{r(1-p)}{p}, V(X) = r(1-p)/p^2$$

ab kuch udaharano par aate hain: pehla example maan lijiye ki 20% of all telephones of a particular model jo warranty ke dauran servicing ke liye aate hain, unme se 60% repair ho sakte hain aur baaki 40% replace kiye jate hain; agar koi company aise 10 telephones kharidti hai to what is the probability ki exactly 2 replace honge under warranty? Pehle hum dekhte hain ki ye kaunsa model fit karta hai har telephone independent hai, har telephone ya to replace hota hai ya nahi hota, aur har trial me probability same hoti hai, to ye binomial experiment hai; define karte hain X = number of telephones replaced under warranty; to X binomial(n=10, p=?); p ka value kya hai? Hume diya gaya tha: 20% telephones servicing ke liye jaate hain, aur unme se 40% replace hote hain; to P(A)=0.2 where A = "submitted for servicing", aur P(B|A)=0.4 where B="replaced"; hume P(B) chahiye: P(B)=P(A and B)=P(B|A)×P(A)+P(B|A^c)=0.4×0.2 + 0 = 0.08; to X ~ binomial(10, 0.08), aur required probability P(X=2)= 10C2 × $(0.08)^2$ × (0.92)^8 = 45 × (0.0064) × (0.5132) ≈ 0.1478; E[X]=np = 10×0.08 = 0.8; Var(X)=np(1–p)=10×0.08×0.92=0.736; standard deviation = √0.736; next example: ek married couple chahte hain ki do betiya ho jaye, maan lete hain ki har birth me boy ya girl ki probability 0.5 hai; what is the probability ki family me exactly 4 children honge? Yani in 4 me se bilkul 2 betiya hain jaise hi doosri beti janam leti hai, births ruk jaati hain; define X = number of boys born before second girl; to ye negative binomial random variable hai jisme r=2 aur p=0.5; hume chahiye P(X=2) kyunki agar X=2 boys, to total children = X + r = 4; PMF se: P(X=2)= (2+2–1 choose 2) × (0.5)^2 × (1–0.5)^2 = (3 choose 2) × (0.5)^4 = 3 × 1/16 = 3/16; expected value E[X] = r(1–p)/p = 2(1–0.5)/0.5 = 2; variance = r(1–p)/p^2 = 2(0.5)/$(0.5)^2$ = 4; expected number of children = E[X+2]=E[X]+2 = 2+2 = 4; aur variance of (X+2) = variance of X = 4; aur R=1 ke special case me ye geometric distribution ban jata hai, yaani geometric random variable: X ~ geometric(p) if X ~ negative binomial(1,p), yani success ki pehli occurrence tak trials karte rehte hain ye case hum pehle bhi dekh chuke hain. directlyHum, to P(X) negative binomial ki PMF to hum pehle hi derive kar chuke hain, to usme r=1 daal denge to directly mil jata hai: P(X)= p(1–p)^x, kyunki agar sirf ek hi success hona hai to success ki probability p hogi aur x failures ki probability (1–p)^x hogi; isme koi binomial coefficient nCr type term nahi aata kyunki yaha success last position par fixed hota hai, isliye ek hi parinaam hai jisme X failures pehle aate hain aur uske baad success hota hai, aur X values hongi 0,1,2,…; isi ko kehte hain geometric random variable with parameter p ye negative binomial ka special case hai; aur expected value

$$E(X) = \frac{r(1-p)}{p}, V(X) = r(1-p)/p^2$$

lekin ek caution: yahan X = number of failures before first success; pehle humne ek aur random variable dekha tha jisme hum count karte the number of trials till first success, to wahan random variable X nahi balki Y define hota hai where Y = number of trials till first success; agar

X failures hai aur last trial me success hai, to Y = X + 1; Y ko bhi kai kitaabon me geometric random variable p ke naam se likha jata hai; Y ki PMF hoti hai: $P(Y=y) = (1-p)^{(y-1)}p$, jahan y=1,2,3,…; expected value of

$$Y = E[X+1] = E[X] + 1 = (1-p)/p + 1 = 1/p;$$

variance of Y = variance of X (kyunki constant add karne se variance nahi badalta) = $(1-p)/p^2$; to dono X aur Y dono ko geometric random variables kaha jata hai, isliye jab kabhi geometric random variable aaye to yeh samajhna zaroori hai ki wo X hai (number of failures before first success) ya Y hai (number of trials till first success) dono related hain par alag random variables hain. Thank you.