Introduction to Probability & Statistics
Prof Abhay Gopal Bhatt
Department of Statistics
Indian Statistical Institute Delhi
Week - 10
Lecture - 34
Estimation

ham dekh rahe the linear combination ka special case: agar hmare paas x1, x2, … xn independent normal random variables hain jinka distribution normal hai lekin iid assume nahi kiya matlab har variable ka mean alag ho sakta hai (mui), variance alag ho sakta hai (sigma_i²), tab koi bhi linear combination jisko hum likhte hain

$$\sum_{i=1}^{n} a_i x_i \sim N(\sum_{i=1}^{n} a_i \mu_i, \sum_{i=1}^{n} a_i{}^2 \sigma_i{}^2)$$

uska exact distribution bhi normal hoga, jiska mean hoga Σ(ai·mui) aur variance hoga Σ(ai²·sigma_i²); ek special corollary hai jab x1 − x2 ko consider karte hain yeh bhi linear combination hi hai jahan a1 = 1, a2 = −1, to distribution normal rahega with mean (mu1 − mu2) aur variance hoga sigma1² + sigma2² (yahaan minus nahi hota — kyunki (−1)² = +1 hota hai); ab ek practical example: petrol pump par teen tarah ke petrol bikte hain jin ki costs hain ₹95/L, ₹100/L, ₹102/L; maan lijiye expected sales volumes (litres) hain E[x1]=1000, E[x2]=500, E[x3]=300 aur variances respectively 100, 80, 50; hum define karte hain y = total revenue = 95x1 + 100x2 + 102x3 — jo ek linear combination hai; to expected value of y = 95·E[x1] + 100·E[x2] + 102·E[x3] = 95000 + 50000 + 30600 = ₹175600; variance of y (independence ke under) = 95²·Var(x1) + 100²·Var(x2) + 102²·Var(x3) + cov terms(=0) = 2222700; to standard deviation = √2222700 ≈ 1490.8720; ab assume kare ki x1, x2, x3 normal distributed bhi hain, to y ka distribution normal hoga N(175600, 2222700); agar hume probability chahiye ki total sale ≥ 172000 — to hum standardize karenge:

$P(y \geq 172000) = P(Z \geq (172000 - 175600) / 1490.8720) = P(Z \geq -2.41) = 1 - \Phi(-2.41) = 0.9920$

(table se); is tarah se agar original xi normal hote hain to hum exact distribution ke saath probability nikal sakte hain; ab yahin par hum probability theory se pure statistics ke core concepts ki taraf transition karte hai jise hum kehte hain estimation ya aakalan jo agle section ka topic hoga. even sampling distributions jab padh rahe the tab hum kya kar rahe the humne alag-alag probability distributions dekhe, jaise Bernoulli(p), Binomial(n,p), Poisson(μ), Normal(μ,σ²), Exponential(λ), Uniform(a,b) etc., aur in sab distributions mein ek ya ek se zyada parameters hote hain ye parameters unknown hote hain; agar koi hume coin de de to hum dekh ke nahi bata sakte ki p kya value hai hume us coin ko toss karke data observe karna padta hai; isi tarah population ko hum unknown-parameter distribution se model karte hain for example coin toss → Bernoulli(p), multiple tosses → Binomial(n,p), rare events → Poisson(μ), lifetimes →

Exponential(λ), tube-light lifetime study etc.; phir population se random sample lete hain: sample $x_1, x_2, \ldots, x_n$; phir sample mean $\bar{x}$ use karke population mean μ ka lagate hain, sample variance se sigma² ka lagate hain ye hi statistical inference hota hai unknown population-parameters ke baare mein conclusion data ke basis par; ab hum enter karte hain estimation concepts: unknown parameter ko hum general symbol θ se denote karte hain; agar hum θ ke liye ek single numeric value estimate karte hain to use kehte hain point estimate; point estimate paane ke liye pehle hum ek suitable statistic choose karte hain (estimator), jaise μ ke liye $\bar{x}$; fir sample-data use karke statistic ki value compute karte hain wo numeric result hota hai point estimate; statistic khud ek random variable hota hai usko kehte hain point estimator ($\hat{\theta}$), aur uski computed numeric value ko kehte hain point estimate; example ke through ye sab clear ho jayega so ab hum examples par aage badhte hain. ye example ek jana-pehchana example hai: maan lo p denote karta hai probability of heads ek coin-toss me; agar x denote kare number of heads in one toss to x sirf 0 ya 1 ho sakta hai, matlab x Bernoulli-p distribution follow karta hai; population se hum random sample lete hain x1, x2, …, x5 sab iid Bernoulli-p; sample mean $\bar{x} = (x1 + \ldots + x5)/5$ represent karta hai proportion of heads in 5 tosses ye hi hum estimator choose karte hain: $\hat{p} = \bar{x}$; lekin point estimate alag-alag samples ke liye alag value deta hai jaise sample1 me agar 100 tosses me 43 heads aaye to $\hat{p} = 0.43$, aur sample2 me 100 tosses me 49 heads aaye to $\hat{p} = 0.49$; isse difference clear hota hai between estimator (random variable $\bar{x}$) aur estimate (actual computed numeric value). doosra example: 5 students ka sample liya gaya to measure physical-exercise time per day; maan lo sample values (minutes) aaye: 75, 105, 60, 30, 50; hume unknown population mean μ estimate karna hai; estimator1: sample mean $\bar{x} = (75+105+60+30+50)/5 = 64$; estimator2: sample median $\tilde{x}$ values ko increasing order me arrange karte hain: 30, 50, 60, 75, 105 beech wali value 60 hai, to $\hat{\mu}2 = 60$; estimator3: (min + max) / 2 = (30 + 105)/2 = 67.5; ye teenon estimators sensible lagte hain lekin inki quality alag hogi; ek hi population mean μ ke liye humne 3 different point estimates paaye: 64, 60, 67.5; ye example dikhata hai ki ek hi estimation problem ke liye multiple possible estimators exist karte hain aur upcoming discussion me hum sikhenge ki kaun-sa estimator “better” hota hai aur kaise choose karna chahiye.