Introduction to Probability & Statistics
Prof Abhay Gopal Bhatt
Department of Statistics
Indian Statistical Institute Delhi
Week - 12
Lecture - 43
Tests of Hypotheses – 2

ham testing of statistical hypothesis dekh rahe the jisme do contradictory claims hote hain ek prior belief hota hai jise hum null hypothesis (H0) kehte hain, aur doosra naya claim hota hai jise hum alternate hypothesis (H1) se denote karte hain; in dono ko equal footing par evaluate nahi kiya jata balki H0 default assumption hota hai (jaise legal system me presumption of innocence), aur H1 sirf tab accept hota hai jab uske support me strong statistical evidence hota hai; test ke liye hum do steps follow karte hain (1) ek suitable test statistic choose karna (jo sample ka function hota hai, jisme unknown parameter nahi hota), jaise mean ke liye x-bar, variance ke liye s^2, etc., aur (2) ek rejection region set karna ek set of values jisme agar test statistic fall kare to hum H0 reject kar denge; lekin kyunki test statistic ek random variable hai, wo chance ke basis par galat decision bhi de sakta hai, isliye testing me errors possible hain: agar H0 actually true ho aur hum ise reject kar dein to type-I error hoti hai, jisko α se denote kiya jata hai (probability of rejecting H0 when H0 is true); doosri taraf, agar H1 true ho aur hum H0 ko reject na karein to type-II error hoti hai, jise β se denote kiya jata hai (probability of not rejecting H0 when H0 is false); ideally hum chahenge ki $\alpha = 0$ aur $\beta = 0$ ho, par real-world me ye impossible hai jab tak hum pura population examine na karein jo practically prohibitively expensive ya impossible hota hai; isliye goal hota hai ki dono errors ki probability chhoti ho, par saath hi hum ek important fact note karte hain: H0 me humesha equality hoti hai ($\theta = \theta_0$), isliye α ek single concrete probability hoti hai kyunki H0 ke under θ ek fixed value hai; par H1 me θ infinite possible values leta hai jaise $\theta > \theta_0$ ya $\theta < \theta_0$ ya $\theta \neq \theta_0$ jiske result me β ek single number nahi hota, ye θ ki value par depend karta hai jo H1-region me kaha padti hai; isi tarah hypothesis testing ka framework systematically type-I error ke control par based hota hai aur H1 ke liye rejection evidence ko measure karta hai.us har ek value ke liye β calculate karne par wo alag hota rahega, kyunki wo har alag-alag θ (ya p) ke liye different probability deta hai jiske corresponding H1 true hota hai; is concept ko hum example se samajhte hain—automobile crash-test case me current bumper design ke saath 25% crashes me koi visible damage nahi hota, aur new bumper design propose kiya gaya hai taaki ye percentage badh sake; parameter p define karte hain as true proportion of crashes with no visible damage under the new design, H0: $p = 0.25$ (current level) aur H1: $p > 0.25$ (improvement claim); sample of $n = 20$ test crashes liya jata hai prototype car par aur x denote karta hai number of crashes jisme koi visible damage nahi aata; $x \sim \text{Binomial}(20, p)$; agar H0 true ho to expected value $E[X] = 20 \times 0.25 = 5$; agar H1 true ho to expectation 5 se zyada hoga; isliye rejection region choose karte hain: reject H0 if $x \geq 8$; ab α = probability of type-1 error = P(reject H0 when H0 is true) = $P(x \geq 8 \mid x \sim \text{Binomial}(20, 0.25))$ = Σ (x=8 to 20) $[20Cx \times (0.25)^x \times (0.75)^{(20-x)}] \approx 0.102$; lekin β = probability of type-2 error = P(do not reject H0 when H0 is false) = $P(x \leq 7 \mid \text{H1 true})$; aur kyunki H1 ke under p koi specific value nahi hai balki koi bhi $p > 0.25$ ho sakta hai, isliye β ek single fixed number nahi hota—wo each possible p ke liye alag hota hai; udaharan ke liye agar actual $p = 0.30$ ho, to $\beta(0.30) = P(x \leq 7 \mid x \sim \text{Binomial}(20, 0.30)) = \Sigma$ (x=0 to 7) [20Cx × (0.30)^x × (0.70)^(20-x)]; is tarah hum dekhte hain ki α ek definite single value hoti hai kyunki H0 ke under parameter ek fixed value leta hai, jabki β ek function hota hai jo parameter ki actual value par depend karta hai jab H1 true hota hai, aur ye statistical testing ke fundamental conceptual framework ka core part hai.0.772 to alpha to chhota tha alpha ki value 0.1 thi lekin β kaafi bada lag raha hai 0.7, matlab 77% chance hai type-2 error ka jab p ki sahi value 0.3 hoti hai; phir humne dekha $\beta(0.5)$ kya hoga, wahi step use karke calculation karne par ye 0.132 aata hai, aur ye clearly dikhta hai ki jaise-jaise p ki actual value H0 ke p=0.25 se door hoti jaati hai, waise-waise β ki value ghatti jaati hai, jo bilkul natural baat hai kyunki not rejecting H0 matlab $x \in$ rejection-region-ka-complement, aur rejecting H0 matlab $x \in$ rejection-region, in dono ki probabilities ka sum 1 hota hai ek hi θ ke liye; agar 0.25 ke time $\alpha=0.102$ hai to usi p=0.25 ke liye $P(x<8) = 1-0.102 \approx 0.90$ hogi; agar p=0.3 lenge to ye 0.9 ke hogi, matlab 0.77 aa rahi hai, aur jaise-jaise p=0.25 se door jayenge β aur ghatti jayegi; ab tak examples me humne R (rejection region) bas designate kar diya thoda arbitrariness thi, lekin sawal yeh uthta hai ke 8 ya usse zyada ko rejection-region kyun rakha 7 ya usse zyada kyun nahi, ya 9 ya usse zyada kyun nahi; yahan logic ye hai ki agar R ko chhota karenge (jaise $R = \{9,10,...,20\}$), to α decrease hoga (type-1 error chance kam hoga), lekin β increase hoga (type-2 error chance badhega), kyunki $\alpha = P(x \in R \mid H0 \text{ true})$ aur $\beta = P(x \notin R \mid H1 \text{ true})$, agar R chhota hoga to R-complement bada hoga, isliye β badhega; matlab hum ek saath α aur β dono ko chhota nahi kar sakte hamesha tradeoff rahega; isliye type-1 error (galat rejection of H0) zyada serious mana jata hai, kyunki hamara statistical testing approach H1 ko accept tabhi karta hai jab H0 ke khilaf heavy sample-evidence ho; isi wajah se hum pehle specify karte hain the maximum tolerable value of α (significance level), jaise $\alpha = 0.1, 0.05, 0.01$; phir hum R is tarah choose karte hain ki P(type-1 error) = α ho ya maximum uske kareeb ho, kyunki uss condition me β minimal possible hota hai given α constraint; isliye test jis α level par banaya jata hai use α-level test bolte hain; ab hum ye sab actual test construction me use karenge aur iss course me focus rahega tests for population means, jisme pehla case hoga: population normal distribution ke saath with known sigma, jahan parameter mu ke testing ki procedure complete detail me samjhe jayegi. to hamesha jaise hamara jo statistical procedure hota hai wo base rahega ek random sample par jo ham population se lete hain $x_1, x_2, \ldots, x_n$ normal with mean mu aur variance sigma-square, mu unknown hota hai aur mu ke baare me jo claims hote hain unko test karne ke liye hum $\bar{x}$ use karte hain kyunki $\bar{x}$ ka distribution exact normal hota hai mean mu aur variance sigma^2/n ke saath; null hypothesis hamesha equality claim hota hai yaha H0: mu = mu0 aur jab H0 true hota hai tab $\bar{x} \sim N(\text{mu0}, \text{sigma}^2/n)$, to hum usko standardize karke Z statistic banate hain: $Z = (\bar{x} - \text{mu0})/(\text{sigma}/\sqrt{n})$, jo standard normal $\sim N(0,1)$ hota hai; agar H1: mu > mu0 test kar rahe hain, to naturally agar actual mu > mu0 hai to hum expect karenge $\bar{x} >$ mu0 lekin agar $\bar{x}$ sirf thoda sa hi mu0 se bada ho to wo strong evidence nahi mana jayega; lekin agar $\bar{x}$ considerable amount se bada ho to Z ki value bhi considerably positive hogi, aur hum Z compute kar sakte hain sample ke $\bar{x}$ se; rejection region tab hoga $Z \geq C$ for some constant $C > 0$, jisme C ko hum alpha (significance level) ke basis par choose karte hain, kyunki alpha = P(type-1 error) = probability of rejecting H0 when H0 is true = P(Z ≥ C | H0 true), aur ye C actually wo hi hota hai jo hum earlier Z_alpha ke naam se define kar chuke hain jiske right-tail area = alpha hota hai, i.e., P(Z > Z_alpha) = alpha; graphically ye standard normal curve hota hai jaha right-side shaded area alpha hota hai; isliye rejection region ban jata hai: reject H0 if Z ≥ Z_alpha aur is test ko upper-tail test kehte hain; isi tarah agar H1: mu < mu0 hota to lower-tail test ban jata aur agar H1: mu ≠ mu0 hota to two-tailed test ye sab hum agle steps me systematically derive karenge.