Introduction to Probability & Statistics
Prof. Abhay Gopal Bhatt
Department of Statistics
Indian Statistical Institute (Delhi)
Week - 2
Lecture - 8
Bayes Theorem

Pichhle class me humne conditional probability ka definition aur examples dekhe the. Jis example par hum kaam kar rahe the, uska sankshipt punaravaartan yeh tha ki ek shop me teen alag-alag brand ke mobile phones bikte hain brand 1 ko 50% log choose karte hain, brand 2 ko 30% aur brand 3 ko 20%. Har phone ke saath ek saal ki warranty milti hai, aur yeh dekha gaya hai ki brand 1 ke 25% phones ko warranty period me repair ki zaroorat padti hai, brand 2 ke 20% ko aur brand 3 ke 10% ko. Humne pichhli baar probability nikali thi ki randomly chosen customer ne brand 1 phone bhi kharida aur usse warranty me repair ki zaroorat bhi padi, jiska answer 0.125 aaya tha. Ab isi example ko aage badha kar tree diagram ke roop me dikhate hain. Tree me pehle teen branches hongi brand 1, brand 2, brand 3 jin par probabilities 0.5, 0.3 aur 0.2 likhi jayengi. Har branch se do branches aur chalengi—repair (B) aur no repair (B complement). Hume conditional probabilities di hui hain: $P(B|A1) = 0.25$, $P(B|A2) = 0.2$, $P(B|A3) = 0.1$. $A1 \cap B$ ki probability nikalne ke liye hum tree me A1 wali branch (0.5) aur uski repair branch (0.25) ko multiply karte hain, jo deta hai 0.125. Isi tarah $A2 \cap B = 0.3 \times 0.2 = 0.06$ aur $A3 \cap B = 0.2 \times 0.1 = 0.02$. Agar hume P(B) chahiye—to ki randomly chosen phone ko warranty me repair ki zaroorat padti hai toh hume inn teenon ko jodna padega: $0.125 + 0.06 + 0.02 = 0.205$. Ab tree ke neeche wale “no repair” branches ke numbers nikalne ke liye conditional complements use karte hain: $1 – 0.25 = 0.75$ (brand 1 ke liye), 0.8 (brand 2 ke liye) aur 0.9 (brand 3 ke liye). Inke intersection probabilities hongi: $A1 \cap B^c = 0.5 \times 0.75 = 0.375$, $A2 \cap B^c = 0.3 \times 0.8 = 0.24$, $A3 \cap B^c = 0.2 \times 0.9 = 0.18$. Inka total 0.795 aata hai aur $P(B)+P(B^c)=1$ hone ki property ke mutabik bilkul sahi baithta hai. Is tarah tree diagram conditional probabilities ko visualize karne ka bahut hi powerful aur aasaan tarika Humne pichhle hisse me conditional probability aur us example ko detail me dekha tha. Ab hum ussi ko aage badha rahe hain. Humne ek customer ko randomly choose kiya aur dekha ki uske phone ko warranty period me repair ki zaroorat padi. Ab hum pooch rahe hain ki is condition ke baad yeh phone brand 1 ka hone ka kya chance hai. Yaani hume $P(A1 \mid B)$ chahiye. Humare paas P(A1), P(A2), P(A3) jaisi prior probabilities hain aur P(B|A1), P(B|A2), P(B|A3) jaisi conditional probabilities bhi. Direct data me P(A1|B) nahi diya, isliye hum conditional probability ka definition use karte hain: $P(A1|B) = P(A1 \cap B) / P(B)$. Upar humne $P(A1 \cap B) = 0.125$ aur $P(B) = 0.205$ calculate kiya tha. Dono ko divide karne par answer aata hai 25/41 jo kaafi high probability hai. Isi tarah hum P(A2|B) aur P(A3|B) bhi nikaal sakte hain: $P(A2|B) = 12/41$ aur $P(A3|B) = 4/41$. Yeh inversion isiliye possible hai kyunki humne Bayes theorem ka framework use kiya. Ab hum general definition likhte hain: Agar A1, A2… An mutually exclusive aur exhaustive events hon, to koi bhi event B ko hum inke intersection ke union ke roop me likh sakte hain: $B = (B \cap A1) \cup (B \cap A2) \dots \cup (B \cap An)$. In sab ka intersection empty hota hai, aur union B ko banata hai. Isse hume Probability ka "Law of Total Probability" milta hai: P(B) = Σ P(Ai) P(B|Ai). Isi se Bayes theorem niklti hai: P(Ai|B) = [P(Ai) P(B|Ai)] / Σ P(Aj) P(B|Aj). Yahan P(Ai) "prior probability" hai yaani shuruaati estimate, aur P(Ai|B) "posterior probability" hai yaani event B ke hone ke baad revised estimate.

$$P\left(\frac{A_i}{B}\right) = \frac{P(A_i \cap B)}{P(B)} = [\mathrm{P(Ai)\ P(B|Ai)}] \,/\, \Sigma\, \mathrm{P(Aj)\ P(B|Aj)}$$

Ab ek bahut important medical example dekhte hain: Ek rare disease population me sirf 0.001 proportion me hoti hai. Ek medical test is disease ko detect karta hai. Agar kisi ko disease hai to test 99% times sahi positive deta hai, aur 1% times false negative. Agar kisi ko disease nahi hai to test fir bhi 2% cases me false positive deta hai. Hume pata karna hai: Agar randomly test kiya gaya vyakti ka test positive aaya, to uske disease hone ki probability kya hai? Yani P(A1 | B). Bayes theorem lagate hain: numerator = 0.001 × 0.99. Denominator = (0.001 × 0.99) + (0.999 × 0.02). Calculation ke baad answer lagbhag 0.047 aata hai, yaani 5% se bhi kam! Yaane test positive aane par bhi vyakti ke bimaar hone ka chance bahut kam hai. Yeh surprising result isliye aata hai kyunki disease bohot rare hai aur false positive rate (2%) rare disease ke context me bahut bada number ban jaata hai. Is example se pata chalta hai ki Bayes theorem real-world medical testing me kitni important hai sirf test positive sun kar result conclude nahi kiya ja sakta, hamesha prior probability ka effect hota hai.