संख्याएँ कहती हैं कहानियाँ

एस. अनंतनारायणन

कुदरती प्रक्रियाओं के कुछ गुणधर्म होते हैं जो तब गड़बड़ हो जाते हैं, जब उनके साथ कोई बाहरी क्रिया की जाती है। 'सामान्यता' के संख्यात्मक द्योतक तब किसी असामान्यता का संकेत दे सकते हैं। ये संकेत ऐसे हो सकते हैं कि उस क्रिया के लिए ज़िम्मेदार अपनी क्रिया को छिपा न पाएँ या डिटेक्शन की सामान्य प्रक्रिया में ज़्यादा समय या प्रयास लगे।

रेल व्हील कारखाना, बैंगलुरु में आयोजित भारतीय रेलवे लेखा सेवा की वार्षिक बैठक में भारतीय प्रबन्धन संस्थान, बैंगलुरु में डिसीज़न साइन्सेज़ एंड इन्फॉर्मेशन सिस्टम्स के सदस्य और गणितज्ञ प्रोफेसर शंकर वेंकटगिरी ने संख्याओं के कुछ गुणधर्मों का विवरण दिया और बताया कि कैसे 'फर्जीवाड़ा पहचान एजेंसियाँ' और कारोबार-जगत संख्याओं के पैटर्न का उपयोग खतरों और अवसरों को पहचानने के लिए करते हैं।

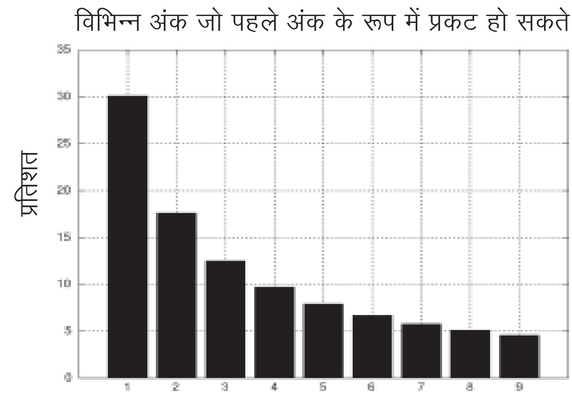
संख्याओं का एक कम जाना-पहचाना गुणधर्म है कि किसी भी कुदरती प्रक्रिया में इन संख्याओं के पहले अंक का वितरण समानता से नहीं होता, बल्कि बड़े अंकों (जैसे 8 व 9) की अपेक्षा छोटे अंकों (1 व 2) की ओर ज़्यादा झुका होता है। उदाहरण के लिए, पर्वतों की ऊँचाइयाँ (फुट में) या इमारतों की ऊँचाइयाँ (मि.मी. में) ऐसी संख्याएँ होंगी जो चन्द सैकड़ा से लेकर कई हज़ार तक होंगी। वास्तविक संख्याओं (12,335 या 8322 या 6345) का पहला अंक 1, 8 और 6 है। क्या इस पहले अंक के 1 से 9 के बीच समान रूप से वितरित होने की बजाय किसी एक रेंज के अन्दर रहने की सम्भावना ज़्यादा होगी?

साधारण रूप से हम उम्मीद करेंगे कि 1 से लेकर 9 तक के सारे अंकों के पहला अंक होने की सम्भावना बराबर

फ्रेंक बेनफोर्ड

होनी चाहिए क्योंकि ये संख्याएँ तो एक बड़े परास (रेंज) में फैली हैं (न कि किसी सीमित परास में)। प्रोफेसर वेंकटगिरी ने बताया कि सहजबोध के विरुद्ध एक नियम है जो कहता है कि ऐसा नहीं है। उन्होंने बताया कि बेनफोर्ड नियम कहता है कि पहले अंक के रूप में 1 पूरे 30 प्रतिशत बार प्रकट होता है जबकि पहले अंक के रूप में 9 मात्र 4.6 प्रतिशत बार। चित्र-1 में दिखाया गया है कि पहले अंक के रूप में विभिन्न अंक कितनी बार प्रकट होते हैं और कैसे यह प्रतिशत 30.1 से 4.6 प्रतिशत तक गिरता है।

इस नियम को कई सारे मामलों में सत्यापित किया गया है कि प्रथम अंक कोई छोटा अंक ज़्यादा बार होता है। जैसे, किसी ज़िले में तालाबों का क्षेत्रफल, विभिन्न गाँवों-कस्बों की जनसंख्याएँ, जन्म या मृत्यु दर, बिजली के बिल, वस्तुओं की कीमतें वगैरह। गौरतलब है कि ये संख्याएँ 'कुदरती' रूप से उभरती हैं। या कह सकते हैं कि ये संख्याएँ इस तरह उभरी हैं कि उसमें प्रथम अंक को प्रभावित करने की कोई योजना नहीं थी। यह बात, उदाहरण के लिए, 12 वर्षीय बच्चों के कद (इंचों में) के बारे में नहीं कही जा सकती क्योंकि वह 50 और 60 इंच के बीच ही होगा। इनमें सबसे ज़्यादा बार प्रथम अंक 5 होगा। दूसरी ओर,



चित्र-1: बेनफोर्ड के नियम का रेखाचित्रीय वर्णन।

किसी तालाब का क्षेत्रफल (वर्ग मीटर में) चन्द सैकड़ा से लेकर हज़ार और शायद सैकड़ों हज़ार वर्ग मीटर तक हो सकता है।

हालाँकि, संख्याओं के प्रथम अंक के बड़े की बजाय छोटे होने का यह गुणधर्म पहली नज़र में थोड़ा अविश्वसनीय या आश्चर्यजनक लगे किन्तु थोड़ा-सा विश्लेषण करके इसे समझा जा सकता है। अंक 1 प्रथम अंक के रूप में सबसे पहले तो संख्या '1' में आएगा और फिर '10-19' तक की संख्याओं में तथा '100-199', '1000-1999' वगैरह में आएगा। इसी प्रकार से प्रथम अंक के रूप में 2 स्वयं संख्या '2' और '2-29', '200-299' वगैरह में आएगा। अंक 3 पहले तो संख्या '3' में और फिर '30-39', '300-399' वगैरह में प्रकट होगा।

आप देख सकते हैं कि अंक '1' सबसे पहले प्रकट होने के बाद मात्र 9 अंकों के बाद फिर दिखने लगता है और उसके बाद 80 संख्याओं के अन्तराल के बाद फिर से दोहराया

बेनफोर्ड का नियम

किसी भी संख्या को प्राप्त करने के लिए 10 को जितनी घात से बढ़ाना पड़ता है, वह उस संख्या का लॉगरिदम कहलाता है। लॉगरिदम का उपयोग यह है कि यदि हमें संख्याओं के लॉग पता हैं, तो हम मात्र उन लॉग्स को जोड़कर संख्याओं का गुणनफल पता कर सकते हैं। इसीलिए सारी संख्याओं के लॉग की गणना 0 से लेकर 1 तक (कई दशमलव अंकों तक) करके तालिकाबद्ध किया गया है। केल्कुलेटर और कम्प्यूटर से पहले ये तालिकाएँ विज्ञान व टेक्नोलॉजी सम्बन्धी कार्यों में सदियों तक काफी उपयोगी रही हैं।

वर्ष 1881 में खगोलविद साइमन न्यूकॉम्ब ने देखा कि लॉगरिदम पुस्तिका के शुरुआती पृष्ठ इस्तेमाल हो-होकर फट चुके थे जबकि बाद के पृष्ठों के साथ ऐसा नहीं था। इसका मतलब यह हुआ कि वैज्ञानिक कार्यों में '1', '2' जैसे छोटे अंकों से शुरू होने वाली संख्याएँ ज़्यादा उभरती हैं, बनिस्बत '8', '9' जैसे बड़े अंकों से शुरू होने वाली। भौतिक शास्त्री फ्रेंक बेनफोर्ड ने यही बात 1938 में फिर से देखी और उन्होंने विभिन्न विषयों में उभरने वाली संख्याओं पर गौर किया। जैसे '335 नदियों की सतह का क्षेत्रफल, यू.एस. की 3359 बस्तियों की जनसंख्या, 104 भौतिक स्थिरांक, 1800 अणु भार, एक गणित हैंडबुक में 5000 प्रविष्टियाँ, रीडर्स डाइजेस्ट के एक अंक में 308 संख्याएँ, 342 अमरीकी विज्ञान पुरुषों के पते, और 418 मृत्यु दरें।' (विकीपीडिया)

इसके बाद बेनफोर्ड ने यह नियम प्रतिपादित किया जिसे उनके नाम से जाना जाता है।

जाता है।

लेकिन अंक ‘2’ को पहली बार दिखने के बाद प्रथम पुनरावृत्ति के लिए 18 संख्याओं का इन्तज़ार करना पड़ता है और उसके बाद 170 संख्याओं का। पुनरावृत्ति का अन्तराल इस प्रकार से अंक बढ़ने के साथ बढ़ता जाता है। अंक ‘9’ को प्रथम पुनरावृत्ति के लिए 90 और दूसरी पुनरावृत्ति के लिए 799 संख्याओं का इन्तज़ार करना पड़ेगा। यह अंक ‘1’ के 80 संख्या के अन्तराल से 10 गुना अधिक है। जैसे-जैसे हम बड़ी संख्याओं की ओर बढ़ते हैं, प्रथम अंक के रूप में उच्चतर अंकों की पुनरावृत्ति का अन्तराल तेज़ी-से (exponentially) बढ़ता है।

कारोबर में अनुप्रयोग

इसी वजह से बड़े परास में फैली संख्याएँ प्रथम अंक के वितरण में बेनफोर्ड नियम का पालन करती हैं। इसका एक इस्तेमाल तो सीधा-सादा है - किसी तंत्र में उत्पन्न संख्याओं में देखना कि क्या उनका प्रथम अंक बेनफोर्ड नियम का पालन करता है। उदाहरण के लिए बैंकों में एक किस्म के फर्जीवाड़े का सम्बन्ध बैंक बेलेंस पर प्रतिदिन ब्याज की गणना से है। जालसाज़ सिस्टम के साथ छेड़छाड़ करके हज़ारों खातों की ब्याज राशि में एक छोटा-सा अंक जोड़ देता है और उसे एक अलग खाते में जमा करता है जिसमें से वह पैसा निकाल सके। यदि बैंक में बेनफोर्ड नियम जाँच मशीन लगी हो तो वह बैंक रिकॉर्ड में नियमित रूप से प्रथम अंक की जाँच करेगी और साथ ही संख्याओं के कुछ अन्य गुणधर्म पर भी नज़र रखेगी। यदि सब कुछ ठीक-ठाक है तो संख्याएँ बेनफोर्ड के नियम का पालन करेंगी। किन्तु यदि कोई व्यवस्थित फेरबदल किया जा रहा है, तो यह बात प्रथम अंक की पुनरावृत्ति में झलक जाएगी और बैंक प्रबन्धन चौकन्ना हो जाएगा।

इसी प्रकार का एक इस्तेमाल सर्वेक्षणों में एकत्रित आँकड़ों के सन्दर्भ में भी हो सकता है। ईमानदारी से किए गए सर्वेक्षण के आँकड़ों के गुणधर्म मनगढ़न्त आँकड़ों या नमूना चुनने में गलतियों से ग्रस्त सर्वेक्षण के आँकड़ों में नज़र नहीं आते। आँकड़ों के सांख्यिकीय विश्लेषण से पता चल जाएगा कि किस तरह के सुधारों की ज़रूरत है। इस तरह की जाँच सांख्यिकीय गुणवत्ता परीक्षण या सुरक्षा सम्बन्धी मामलों में निहायत ज़रूरी होती है।

एस. अनंतनारायणन: विज्ञान लेखक, सलाहकार, मध्यस्थ (औद्योगिक विवाद), वकील।
अँग्रेज़ी से अनुवाद: सुशील जोशी: एकलव्य द्वारा संचालित *स्रोत* फीचर सेवा से जुड़े हैं। विज्ञान शिक्षण व लेखन में गहरी रुचि।
यह लेख *न्यू साइंटिस्ट* पत्रिका के अंक- नवम्बर 2016 से साभार।