FYBCOM Sem II

Unit 4.3

Hypothesis- Kinds of Hypothesis Testing, Significance of Hypothetical Testing.

પૂર્વધારણા- પૂર્વધારણા પરીક્ષણના પ્રકાર, અનુમાનિત પરીક્ષણનું મહત્વ.

આંકડાશાસ્ત્રમાં પૂર્વધારણા પરીક્ષણ શું છે?

પૂર્વધારણા પરીક્ષણ એ આંકડાકીય વિશ્લેષણનો એક પ્રકાર છે જેમાં તમે વસ્તી પરિમાણ વિશેની તમારી ધારણાઓને પરીક્ષણમાં મૂકો છો. તેનો ઉપયોગ 2 આંકડાકીય ચલો વચ્ચેના સંબંધનો અંદાજ કાઢવા માટે થાય છે.

ચાલો વાસ્તવિક જીવનમાંથી આંકડાકીય પૂર્વધારણાના થોડા ઉદાહરણોની ચર્ચા કરીએ -

- એક શિક્ષક ધારે છે કે તેની કૉલેજના 60% વિદ્યાર્થીઓ નિમ્ન-મધ્યમ-વર્ગના પરિવારોમાંથી આવે છે.
- ડૉક્ટર માને છે કે ડાયાબિટીસના દર્દીઓ માટે 3D (આહાર, માત્રા અને શિસ્ત) 90% અસરકારક છે.

હવે જ્યારે તમે પૂર્વધારણા પરીક્ષણ વિશે જાણો છો, ત્યારે આંકડાઓમાં બે પ્રકારના પૂર્વધારણા પરીક્ષણને જુઓ.

આજના ડેટા-આધારિત વિશ્વમાં, નિર્ણયો હંમેશા ડેટા પર આધારિત હોય છે. પૂર્વધારણા તે પ્રક્રિયામાં નિર્ણાયક ભૂમિકા ભજવે છે, પછી ભલે તે વ્યવસાયિક નિર્ણયો લેતા હોય, આરોગ્ય ક્ષેત્રે, શિક્ષણમાં અથવા ગુણવત્તા સુધારણામાં. પૂર્વધારણા અને પૂર્વધારણા પરીક્ષણો વિના, તમે ખોટા તારણો દોરવાનું અને ખરાબ નિર્ણયો લેવાનું જોખમ લે છે. આ ટ્યુટોરીયલમાં, તમે આંકડાશાસ્ત્રમાં પૂર્વધારણા પરીક્ષણ જોશો.

આંકડાશાસ્ત્રમાં પૂર્વધારણા પરીક્ષણ શું છે?

પૂર્વધારણા પરીક્ષણ એ આંકડાકીય વિશ્લેષણનો એક પ્રકાર છે જેમાં તમે વસ્તી પરિમાણ વિશેની તમારી ધારણાઓને પરીક્ષણમાં મૂકો છો. તેનો ઉપયોગ 2 આંકડાકીય ચલો વચ્ચેના સંબંધનો અંદાજ કાઢવા માટે થાય છે.

ચાલો વાસ્તવિક જીવનમાંથી આંકડાકીય પૂર્વધારણાના થોડા ઉદાહરણોની ચર્ચા કરીએ -

- એક શિક્ષક ધારે છે કે તેની કૉલેજના 60% વિદ્યાર્થીઓ નિમ્ન-મધ્યમ-વર્ગના પરિવારોમાંથી આવે છે.
- ડૉક્ટર માને છે કે ડાયાબિટીસના દર્દીઓ માટે 3D (આહાર, માત્રા અને શિસ્ત) 90% અસરકારક છે.

હવે જ્યારે તમે પૂર્વધારણા પરીક્ષણ વિશે જાણો છો, ત્યારે આંકડાઓમાં બે પ્રકારના પૂર્વધારણા પરીક્ષણને જુઓ.

નલ પૂર્વધારણા અને વૈકલ્પિક પૂર્વધારણા

નલ પૂર્વધારણા એ ધારણા છે કે ઘટના બનશે નહીં. એક શૂન્ય પૂર્વધારણાનો અભ્યાસના પરિણામ પર કોઈ અસર નથી સિવાય કે તેને નકારવામાં આવે.

H0 એ તેના માટેનું પ્રતીક છે, અને તેનો ઉચ્ચાર H-નૉટ છે.

વૈકલ્પિક પૂર્વધારણા એ નલ પૂર્વધારણાની તાર્કિક વિરુદ્ધ છે. વૈકલ્પિક પૂર્વધારણાની સ્વીકૃતિ શૂન્ય પૂર્વધારણાના અસ્વીકારને અનુસરે છે. H1 તેના માટે પ્રતીક છે.

આ વાતને એક ઉદાહરણથી સમજીએ.

એક સેનિટાઈઝર ઉત્પાદક દાવો કરે છે કે તેનું ઉત્પાદન સરેરાશ 95 ટકા જંતુઓને મારી નાખે છે.

આ કંપનીના દાવાની ચકાસણી કરવા માટે, એક શૂન્ય અને વૈકલ્પિક પૂર્વધારણા બનાવો.

H0 (નલ પૂર્વધારણા): સરેરાશ = 95%.

વૈકલ્પિક પૂર્વધારણા (H1): સરેરાશ 95% કરતા ઓછી છે.

આ ખ્યાલને સમજવા માટેનું બીજું સીધું ઉદાહરણ એ નક્કી કરી રહ્યું છે કે સિક્કો વાજબી અને સંતુલિત છે કે નહીં. નલ પૂર્વધારણા જણાવે છે કે માથાના પ્રદર્શનની સંભાવના પૂંછડીઓના પ્રદર્શનની સંભાવના જેટલી છે. તેનાથી વિપરીત, વૈકલ્પિક સિદ્ધાંત જણાવે છે કે માથા અને પૂંછડીઓના પ્રદર્શનની સંભાવના ઘણી અલગ હશે.

સરળ અને સંયુક્ત પૂર્વધારણા પરીક્ષણ

વસ્તીના વિતરણના આધારે, તમે આંકડાકીય પૂર્વધારણાને બે પ્રકારમાં વર્ગીકૃત કરી શકો છો.

સરળ પૂર્વધારણા: એક સરળ પૂર્વધારણા પરિમાણ માટે ચોક્કસ મૂલ્યનો ઉલ્લેખ કરે છે.

સંયુક્ત પૂર્વધારણા: સંયુક્ત પૂર્વધારણા મૂલ્યોની શ્રેણીને સ્પષ્ટ કરે છે.

ઉદાહરણ:

એક કંપની દાવો કરી રહી છે કે આ ક્વાર્ટરમાં તેમનું સરેરાશ વેચાણ 1000 યુનિટ છે. આ એક સરળ પૂર્વધારણાનું ઉદાહરણ છે.

ધારો કે કંપની દાવો કરે છે કે વેચાણ 900 થી 1000 યુનિટની રેન્જમાં છે. પછી આ એક સંયુક્ત પૂર્વધારણાનો કેસ છે.

એક પૂંછડી અને બે પૂંછડીની પૂર્વધારણા પરીક્ષણ

વન-ટેઈલ્ડ ટેસ્ટ, જેને ડાયરેક્શનલ ટેસ્ટ પણ કહેવાય છે, તે ડેટાના નિર્ણાયક ક્ષેત્રને ધ્યાનમાં લે છે જેના પરિણામે જો ટેસ્ટનો નમૂનો તેમાં આવે તો શૂન્ય પૂર્વધારણાને નકારી કાઢવામાં આવશે, અનિવાર્યપણે વૈકલ્પિક પૂર્વધારણાની સ્વીકૃતિનો અર્થ થાય છે.

એક-પૂંછડીવાળા પરીક્ષણમાં, નિર્ણાયક વિતરણ ક્ષેત્ર એકતરફી હોય છે, એટલે કે પરીક્ષણનો નમૂનો ચોક્કસ મૂલ્ય કરતાં વધુ અથવા ઓછો હોય છે.

બે પૂંછડીઓમાં, પરીક્ષણ નમૂનાને દ્વિ-પૂંછડીવાળા પરીક્ષણમાં મૂલ્યોની શ્રેણી કરતાં વધુ અથવા ઓછા હોવાનું ચકાસવામાં આવે છે, જે સૂચવે છે કે નિર્ણાયક વિતરણ ક્ષેત્ર બે બાજુ છે.

જો નમૂના આ શ્રેણીમાં આવે છે, તો વૈકલ્પિક પૂર્વધારણા સ્વીકારવામાં આવશે, અને શૂન્ય પૂર્વધારણાને નકારવામાં આવશે.

ઉદાહરણ:

ધારો કે H0: સરેરાશ = 50 અને H1: સરેરાશ 50 ના બરાબર છે

H1 મુજબ, સરેરાશ 50 કરતા વધારે અથવા ઓછો હોઈ શકે છે. આ ટુ-ટેલ્ડ ટેસ્ટનું ઉદાહરણ છે.

તેવી જ રીતે, જો H0: સરેરાશ >=50, તો H1: સરેરાશ <50

અહીં સરેરાશ 50 કરતા ઓછો છે. તેને વન-ટેલ્ડ ટેસ્ટ કહેવામાં આવે છે.

પ્રકાર 1 અને પ્રકાર 2 ભૂલ

પૂર્વધારણા પરીક્ષણ બે પ્રકારની ભૂલોમાં પરિણમી શકે છે.

પ્રકાર 1 ભૂલ: એક પ્રકાર-1 ભૂલ ત્યારે થાય છે જ્યારે નમૂનાના પરિણામો સાચા હોવા છતાં નલ પૂર્વધારણાને નકારે છે.

ટાઈપ 2 ભૂલ: ટાઈપ-I ભૂલથી વિપરીત જ્યારે નલ પૂર્વધારણા ખોટી હોય ત્યારે તેને નકારવામાં ન આવે ત્યારે ટાઈપ-II ભૂલ થાય છે.

ઉદાહરણ:

ધારો કે વિદ્યાર્થી પાસ થાય કે નાપાસ થાય તે નક્કી કરવા શિક્ષક પરીક્ષાના પેપરનું મૂલ્યાંકન કરે છે.

H0: વિદ્યાર્થી પાસ થયો છે

H1: વિદ્યાર્થી નાપાસ થયો છે

Type I ભૂલ શિક્ષક વિદ્યાર્થીને નાપાસ કરશે [H0 નકારે છે] જો કે વિદ્યાર્થીએ પાસિંગ માર્કસ મેળવ્યા હતા [H0 સાચું હતું].

પ્રકાર II ભૂલ એ કેસ હશે જ્યાં શિક્ષક વિદ્યાર્થીને પાસ કરે છે [H0 નકારશો નહીં] જોકે વિદ્યાર્થીએ પાસિંગ માર્કસ મેળવ્યા નથી [H1 સાચું છે].

મહત્વનું સ્તર

આલ્ફા મૂલ્ય એ નક્કી કરવા માટેનો માપદંડ છે કે શું પરીક્ષણ આંકડા આંકડાકીય રીતે મહત્વપૂર્ણ છે. આંકડાકીય પરીક્ષણમાં, આલ્ફા પ્રકાર I ભૂલની સ્વીકાર્ય સંભાવના દર્શાવે છે. કારણ કે આલ્ફા એક સંભાવના છે, તે 0 અને 1 ની વચ્ચે ગમે ત્યાં હોઈ શકે છે. વ્યવહારમાં, સૌથી સામાન્ય રીતે ઉપયોગમાં લેવાતા આલ્ફા મૂલ્યો 0.01, 0.05 અને 0.1 છે, જે 1%, 5%, અને 10% પ્રકાર I ભૂલની સંભાવના દર્શાવે છે. , અનુક્રમે (એટલે કે શૂન્ય પૂર્વધારણાને નકારી કાઢવી જ્યારે તે હકીકતમાં સાચી હોય).

પી-વેલ્યુ

p-મૂલ્ય એ એક મેટ્રિક છે જે સંભાવના વ્યક્ત કરે છે કે અવલોકન કરેલ તફાવત તક દ્વારા આવી શકે છે. જેમ જેમ p-મૂલ્ય ઘટે છે તેમ અવલોકન કરેલ તફાવતનું આંકડાકીય મહત્વ વધે છે. જો p-મૂલ્ય ખૂબ ઓછું હોય, તો તમે નલ પૂર્વધારણાને નકારી કાઢો છો.

અહીં તમે એક ઉદાહરણ લીધું છે જેમાં તમે ચકાસવાનો પ્રયાસ કરી રહ્યા છો કે શું નવી જાહેરાત ઝુંબેશથી ઉત્પાદનના વેચાણમાં વધારો થયો છે. p-વેલ્યુ એ એવી સંભાવના છે કે નલ પૂર્વધારણા, જે જણાવે છે કે નવી જાહેરાત ઝુંબેશને કારણે વેચાણમાં કોઈ ફેરફાર થયો નથી, તે સાચું છે. જો p-મૂલ્ય .30 છે, તો ઉત્પાદનના વેચાણમાં કોઈ વધારો કે ઘટાડો ન થવાની 30% શક્યતા છે. જો p-વેલ્યુ 0.03 છે, તો 3% સંભાવના છે કે નવી જાહેરાત ઝુંબેશને કારણે વેચાણ મૂલ્યમાં કોઈ વધારો અથવા ઘટાડો થયો નથી. જેમ તમે જોઈ શકો છો, પી-વેલ્યુ જેટલું ઓછું હશે, વૈકલ્પિક પૂર્વધારણા સાચી હોવાની શક્યતાઓ વધી જાય છે, જેનો અર્થ છે કે નવી જાહેરાત ઝુંબેશ વેચાણમાં વધારો અથવા ઘટાડોનું કારણ બને છે.

નિષ્કર્ષ

આ ટ્યુટોરીયલ વાંચ્યા પછી, તમને પૂર્વધારણા પરીક્ષણની વધુ સારી સમજ હશે, જે ડેટા સાયન્સના ક્ષેત્રમાં સૌથી મહત્વપૂર્ણ ખ્યાલોમાંની એક છે. મોટાભાગની પૂર્વધારણાઓ અવલોકન કરેલ વર્તન, કુદરતી ઘટનાઓ અથવા સ્થાપિત સિદ્ધાંતો વિશેની અટકળો પર આધારિત છે.