

P-värden och hur man tolkar dem

Peter M. Dahlgren, JMG

2017-03-18

P-värdet kan vara det svåraste att förstå. Här gör jag några pedagogiska försök att skapa en intuitiv förståelse för p-värdet.

Definition av p-värden

P-värdet visar sannolikheten att få ett resultat som är så extremt (eller mer extremt), givet att nollhypotesen är sann.

P-värden kommer från ett statistiskt test som man genomför på en datamängd. P-värdet går från 0 till 1. Ju närmare 0 det är, desto mindre sannolikhet att du har fått resultatet ifråga, givet att nollhypotesen är sann.

Om du får 0,000 i ditt statistikprogram beror det på avrundning. Skriv då $p < 0,001$ med tre decimaler.

Hur man tolkar p-värdet

- Ett **lågt p-värde** (exempelvis $p = 0,001$) innebär att det är osannolikt att få det resultatet du fått, givet att nollhypotesen är sann. Det går därmed säga att det observerade sambandet skiljer sig från ett nollsamband. Annorlunda uttryckt: det går skilja signalen från bruset, och det innebär att ett samband kunde hittas.
- Ett **högt p-värde** (exempelvis $p = 0,583$) innebär att det är sannolikt att få det resultatet du fått, givet att nollhypotesen är sann. Det går därmed inte säga att det observerade sambandet skiljer sig från ett nollsamband. Annorlunda uttryckt: det går inte skilja signalen från bruset, och det innebär att något samband inte kunde hittas.

Ett p-värde på 0,012 betyder exempelvis att det är 1,2 % sannolikhet att få det resultat du fått, eller mer extremt, givet att nollhypotesen är sann. Med andra ord betyder det att det är väldigt osannolikt att sambandet kunde uppstå om det faktiskt inte fanns i populationen. När du tittat på p-värdet tittar du därefter på sambandsmättet för att se hur stort sambandet är.

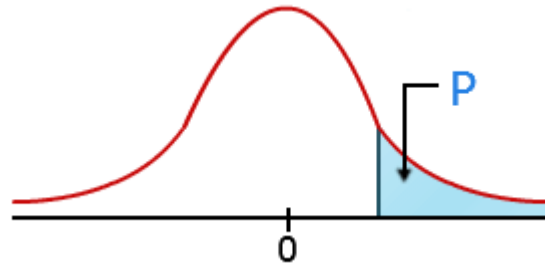


Figure 1: Om p-värdet är lågt, kommer det att vara långt ut till höger eller vänster. Om p-värdet är högt, kommer det att vara runt mitten.

Exempel: placera bollar i korgar



Figure 2: Om man har lika många bollar i båda korgarna är p-värdet högt, eftersom fördelningen är sannolik givet slumpen.

Vi kan ta ett exempel på hur p-värdet fungerar.

1. Du har 100 stycken bollar som ska placeras i två korgar.
2. Om bollarna placeras helt slumpmässigt i korgarna kommer det att vara ungefär 50 bollar i ena korgen och 50 i den andra (50/50). Då säger vi att det inte finns något samband mellan korgarna och bollarna. Detta är nollhypotesen.
3. Låt säga att vi i stället lägger 30 bollar i ena korgen och 70 i den andra

korgen. Detta är den alternativa hypotesen.

4. Vi kan nu jämföra fördelningen 50/50 (nollhypotesen) med fördelningen 30/70 (alternativa hypotesen).
5. Vi kan uttrycka sannolikheten att vi fått detta resultat (30/70) eller mer extremt, givet att bollarna är fördelade slumpmässigt 50/50. Denna sannolikhet är p-värdet.
6. När bollarna placeras 30/70 så är det resultatet så pass extremt, givet fördelningen 50/50, att p-värdet blir väldigt litet (se nedan). Vi förkastar då nollhypotesen (50/50) och accepterar den alternativa hypotesen.
7. Men om bollarna hade placeras 51/49 så är resultatet inte alls extremt, givet fördelningen 50/50, och p-värdet blir då väldigt stort (se nedan). Vi kan då inte förkasta nollhypotesen (50/50).

Det enda p-värdet gör är alltså att se om bollarna har placerats slumpmässigt i korgarna, eller inte. I vetenskap hjälper p-värdet att se om ett samband kan ha uppstått slumpmässigt eller inte.

Här är några exempel på vad p-värdet blir när man flyttar runt bollarna i korgarna:

- Om bollarna är fördelade 50/50 blir $p = 1,0$. Tolkning: Det observerade resultatet är väldigt sannolikt givet slumpen. Det observerade resultatet (50/50) är inte extremt i förhållande till slumpen (50/50) eftersom de är identiska.
- Om bollarna är fördelade 51/49 blir $p = 0,9204$. Tolkning: Det är väldigt sannolikt givet slumpen.
- Om bollarna är fördelade 55/45 blir $p = 0,3682$. Tolkning: Det är fortfarande sannolikt givet slumpen (även om det är något mindre sannolikt).
- Om bollarna är fördelade 30/70 blir $p = 0,0000785$. Tolkning: Det är väldigt osannolikt givet slumpen.
- Om bollarna är fördelade 10/90 blir $p = 0,000000000000000022$. Tolkning: Det är extremt osannolikt givet slumpen.

Vad som påverkar p-värdet

Det är sampelstorleken och effektstorleken som påverkar p-värdet. Det innebär att man får mindre p-värden om man har många analysenheter. I en surveyundersökning är det därför lättare att få lägre p-värden om man frågar fler personer. Det är också lättare att få låga p-värden ju större sambandet är.

I exemplet ovan med bollarna innebär det att fördelningen av bollarna påverkar p-värdet.

Vilka missförstånd om p-värden

Här är några vanliga missförstånd kring p-värden:

- P-värdet visar *inte* sannolikheten att nollhypotesen är sann. (Utgångspunkten är alltid att nollhypotesen är sann.)
- P-värdet visar *inte* vad som är relevant, viktigt eller vetenskapligt intressant. (P-värdet har med slumpen att göra.)
- P-värdet visar *inte* sannolikheten att resultatet beror på slumpen. (Slumpen är utgångspunkten, se första punkten.)
- P-värdet visar *inte* sannolikheten att få samma resultat om gör analysen en gång till.
- P-värdet visar *inte* storleken på sambandet. (Eftersom sampelstorleken skiljer sig åt kommer p-värdet skilja sig åt.)

P-värdet bör inte betraktas svartvit. Ett p-värde på 0,049 är inte bättre än ett p-värde på 0,051. Det är snarare en skala, där lägre p-värde kan göra oss mer säkra på att ett samband finns.

Bli inte vemodig om du har svårt att förstå p-värden. Till och med forskare brottas med frågan: Not Even Scientists Can Easily Explain P-values. Det är ett krångligt koncept som är lätt att få om bakfoten. P-värden har kritiserats i nästan hundra år, kanske starkast av psykologen Paul Meehl som sa att det är “one of the worst things that ever happened in the history of psychology” (Meehl, 1978, s. 817).

Statistisk signifikans

Statistisk signifikans är när p-värden är under en specifik gräns.

Av tradition brukar man säga att p-värden under gränsen 0,05 är *statistiskt signifikanta*, medan p-värden över 0,05 ej är statistiskt signifikanta. Därför ser man ofta $p < 0,05$ i vetenskapliga texter.

Gränsen 0,05 är fullständigt godtycklig och används av historiska skäl. Gränsen brukar kallas alfanivå eller signifikansnivå, och den bestämmer man själv *innan* man gör analysen.

Vad är nollhypotes och alternativ hypotes?

Nollhypotesen är antagandet att det *inte* finns något samband mellan två variabler.

Den alternativa hypotesen är antagandet att det *finns* en skillnad mellan två variabler.

Om du signifikanstestar ett sambandsmått är nollhypotesen detsamma som nollsamband. Med andra ord att Pearsons $r = 0$. Ju längre ifrån 0 sambandet är, desto lägre kommer p-värdet att bli. Statistiska tester utgår alltid från nollhypotesen och testar hur extremt resultatet blir.

P-värdet i SPSS

Titta i kolumnen *Approximate Significance (Approx. Sig.)* eller bara *Sig.*

Rapportera p-värden

P-värden rapporteras alltid tillsammans med ett statistiskt test, till exempel sambandsmått. Skriv därför p-värdet efter sambandsmåtten. Rapportera alltid p-värdet med 3 decimaler.

Exempel med Pearsons r :

Det fanns en medelstor positiv korrelation mellan tidningsläsning och politisk kunskap ($r = 0,31$; $p = 0,002$).

Här syftar r på Pearsons r . Notera att texten i princip skrivs i följande format: Det fanns en [liten/medelstor/stor] korrelation mellan [X] och [Y] ($r = 0,000$; $p = 0,000$).

Du kan inte skriva att tidningsläsning *påverkar* politisk kunskap eftersom vi inte vet den kausala riktningen (om du inte genomför experiment eller dylikt).

Exempel med Cramérs V :

Det fanns en medelstor korrelation mellan tidningsläsning och politisk kunskap (Cramérs $v = 0,31$; $p = 0,002$).

Notera att Cramérs V går från 0 till 1 och därmed inte har någon riktning. Det finns alltså ingen positiv eller negativ korrelation när man har Cramérs V .

Video

Anders Sundell (Statsvetenskapliga institutionen) förklarar signifikans och p-värden i denna video.

Fler videor om p-värden:

- What is a p-value? (6 min)
- Hypothesis testing and p-values - Khan Academy (12 min)

Referenser

- Meehl, P.E. (1978). Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. *Journal of Consulting and Clinical Psychology* 46, 806–834.