Um kenningaprófanir og baysískar ályktanir.

Helgi Tómasson
helgito@hi.is

Bayesistum leyfist að setja fram líkur á að tiltekin kenning sér sönn. „A prioro” líkur á að H_0 sé sönn eru π_0, og valkenning H_1, $\pi_1 = 1 - \pi_0$. Ef bornar eru saman tvær einfaldar kenningar, H_0: rétt líkan er $f_0(x)$ og H_1 rétt líkan er $f_1(x)$. Síðan er gögnum, x, safnað og „a posteriori” er því:

$$P(H_0|x) = \frac{\pi_0 f_0(x)}{\pi_0 f_0(x) + \pi_1 f_1(x)}.$$

Hlutfallið:

$$\frac{P(H_0|x)}{P(H_0|x)} = \underbrace{\frac{\pi_0}{\pi_1}}_{A} \underbrace{\frac{f_0(x)}{f_1(x)}}_{B},$$

lýsir „a posteriori” hlutfalli á líkum að kenningarnar séu sannar. Hluti A er „a priori” hlutfall og hluti B er nefndur „Bayes factor”(BF). BF lýsir því hvernig gögnin, x hafa breytt tiltrú á H_0. Ef BF er stærri en 1 þá er það túlkað sem að líkur á að H_0 sé sönn hafi aukist við gagnasöfnunina. Á hliðstæðan hátt er lítið gildi á BF túlkað sem að gagnasöfnun hafi dregið úr líkum á að H_0 sé sönn.

Ef kenningarnar eru samsettar, t.d. að H_i leyfi fjölskyldu af dreifingum, $H_i : \pi_i(\theta), \quad \theta \in \Theta_i$, þá er

$$\mathrm{BF} = \frac{\int_{\Theta_0} f(x|\theta) d\theta}{\int_{\Theta_1} f(x|\theta) d\theta}.$$

Valið á formi upplýsinga π_i miðað við kenningarnar H_i skiptir hér máli og túlkun getur verið á skjön við það sem tíðkast. Ályktun er háð $\pi_0(\theta)$ og $\pi_1(\theta)$ á máta sem ekki er augljós. Eftirfarandi dæmi er úr kennslbók eftir Young & Smith (2005).

Ef $X_1, \ldots, X_n$ eru óháðar Bernoulli(θ), þ.e. $X = \sum_{i=1}^{n} X_i$ er binomial(n, θ). $H_0 : \theta = \theta_0 = 1/2$ er sett fram og prófu gegn keningunni $H_1 : \theta \quad U(0, 1)$. Með smáalgebru má sjá að BF er:

$$\mathrm{BF} = \frac{(n+1)!}{x!(n-x)!} \theta_0^x (1 - \theta_0)^{n-x}.$$

Ef mælingin er $x = 60$ og $n = 100$ þá fæst að að z-gildi er $z = 2$, þ.e. að hafna ætti H_0 í hefðbundnum skilingi miðað við $\alpha = 0.05$ og álykta að θ væri marktækt frábruðið 1/2. Hægt er að námunda BF með jöfnu Stirling ($n! \simeq \sqrt{2\pi n}(n/e)^n$) má fá að BF er stærri en 1. Þ.e. að tilraunin eyku líkur á að H_0 sé sönn. Í þessu tilfelli gefa því hefðbundin kenningapróf og bayesísk kenningaprófun sitt hvora ályktunina.

Heimildir

Young, G. & Smith, R. (2005). *Essentials of Statistical Inference.* Cambridge University Press.