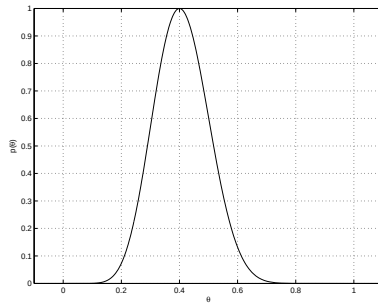


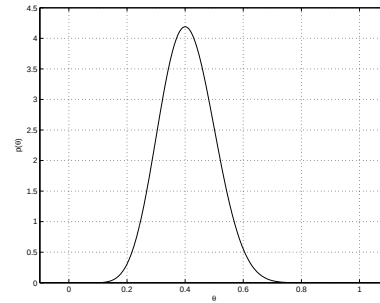
$$\theta = 0.4$$

$$\theta = 0.4 \pm 0.2 (\approx 95\% \text{ konfidens})$$

Exempel på trolighetsfunktion för en sannolikhet θ



Normerad trolighetsfunktion för sannolikheten θ



Definition 1 Vi ska kalla $p(\theta)$ för θ :s *à priori-täthet* eller *à priori-trolighet*.

À priori är latin och betyder på förhand

À priori-tätheten beskriver den kunskap om θ som vi har nu

Vi ska definiera ett försök och en modell $p(x|\theta)$

Obs att mätresultatet x kan vara flerdimensionellt

Exempel 5 Låt θ vara sannolikheten att en viss händelse A inträffar i ett försök och antag att vi gör n oberoende upprepningar av försöket. Låt x vara antalet lyckade försök. Då är $x|\theta$ binomialfördelad med parametrar n och θ . Motsvarande massfunktion är

$$p(x|\theta) = \binom{n}{x} \theta^x (1 - \theta)^{n-x}$$

där $x = 0, 1, \dots, n$.

Exempel 6 Låt $\theta = 1/\lambda$ vara väntevärdet av en exponentialfördelad tid och vi ska göra n oberoende mätningar av denna tid. Låt

$$x = x_1, \dots, x_n$$

beteckna de erhållna mätresultaten. Den täthet som beskriver detta försök är

$$p(x|\theta) = \prod_{i=1}^n \left(\frac{1}{\theta} e^{-x_i/\theta} \right) = \frac{1}{\theta^n} e^{-\sum_i x_i/\theta} = \frac{1}{\theta^n} e^{-n\bar{x}/\theta}$$

där $x_i > 0$ för $i = 1, \dots, n$.

Vi känner

- à priori-tätheten $p(\theta)$,
- mät- eller observationsmodellen $p(x|\theta)$ och
- data x

Vi vill beräkna $p(\theta|x)$

Definition 2 Vi ska kalla $p(\theta|x)$ för θ :s *à posteriori-täthet* eller *à posteriori-trolighet*.

À posteriori är latin och betyder i efterhand

À posteriori-tätheten beskriver vår uppdaterade kunskap om θ som vi har efter det att vi gjort och analyserat försöket

Sats 1 (Bayes sats) Å posterioritätheten $p(\theta|x)$ är proportionell mot produkten av modell-tätheten (eller massfunktionen) $p(x|\theta)$ och å priori-tätheten $p(\theta)$, d v s

$$p(\theta|x) \propto p(x|\theta)p(\theta)$$

Tecknet $\propto$ betyder proportionell mot

Således säger Bayes sats att

$$p(\theta|x) = C(x)p(x|\theta)p(\theta)$$

I fallet då "prior" $p(\theta)$ är en täthet, gäller att

$$C(x) = \left(\int_{\mathcal{R}} p(x|\theta)p(\theta) d\theta \right)^{-1}$$

Då är även "posterior"

$$p(\theta|x) = \frac{p(x|\theta)p(\theta)}{\int_{\mathcal{R}} p(x|\theta)p(\theta) d\theta}$$

en täthet.

Tillämpning 1 (J.f.r exempel 5) Vi är intresserade av sannolikheten θ att en viss händelse A inträffar då vi gör ett slumpmässigt försök. Vi har ingen förhandskunskap om θ och ansätter därför

$$p(\theta) = 1$$

för $0 < \theta < 1$.

Vi gör $n_1 = 7$ oberoende upprepningar av försöket, och erhåller $x_1 = 4$ (händelsen A inträffade alltså 4 gånger).

Modell för hela försöket: $x|\theta$ är $\text{bin}(n, \theta)$ -fördelad med $n = n_1 = 7$. Således är

$$p(x_1|\theta) = \binom{n_1}{x_1} \theta^{x_1} (1-\theta)^{n-x_1}$$

där $x_1 = 0, 1, \dots, n_1$ och $n_1 = 7$.

Bayes formel ger

$$p(\theta|n_1, x_1) \propto \theta^{x_1} (1-\theta)^{n_1-x_1}$$

Efter insättning av observerade data, har vi

$$p(\theta|n_1 = 7, x_1 = 4) \propto \theta^4 (1-\theta)^3$$

Antag nu att vi avser att göra ytterligare $n_2 = 10$ oberoende försök. Funktionen

$$p(\theta|n_1 = 7, x_1 = 4) \propto \theta^4 (1-\theta)^3$$

reflekterar vår nuvarande kunskap om θ . Detta är vår nya prior.

Modell för de nya försöken är

$$p(x_2|\theta) = \binom{n_2}{x_2} \theta^{x_2} (1-\theta)^{n_2-x_2}$$

där $x_2 = 0, 1, \dots, n_2$ och $n_2 = 10$. Bayes sats ger att

$$p(\theta|n_1 = 7, x_1 = 4, n_2, x_2) \propto \theta^{4+x_2} (1-\theta)^{3+n_2-x_2}$$

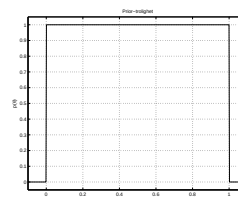
Antag att A inträffade i $x_2 = 6$ försök. Efter insättning har vi slutligen posteriorn

$$p(\theta|n_1 = 7, x_1 = 4, n_2 = 10, x_2 = 6) \propto \theta^{10} (1-\theta)^7$$

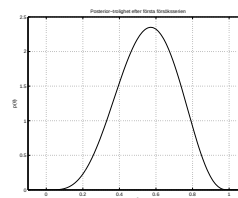
Notera att det är bara totala antalet försök och totala antalet gånger A inträffat, som finns med på högersidan ovan. Alltså kan vi lika gärna skriva

$$p(\theta|n = 17, x = 10) \propto \theta^{10} (1-\theta)^7$$

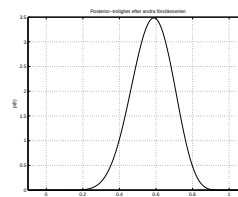
för vår posterior-täthet.



$$p(\theta) = 1$$



$$p(\theta) \propto \theta^4 (1-\theta)^3$$



$$p(\theta) \propto \theta^{10} (1-\theta)^7$$

Beta-fördelningen

Beta(p, q)-fördelningen definieras av att tätheten är

$$f(\theta) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} \theta^{p-1} (1-\theta)^{q-1}$$

för $0 < \theta < 1$.

Beta(p, q)-fördelningens väntevärde, varians och mod är

$$\begin{aligned} E[\theta] &= \frac{p}{p+q} \\ \text{Var}[\theta] &= \frac{pq}{(p+q)^2(p+q+1)} \\ \hat{\theta} &= \frac{p-1}{p+q-2} \end{aligned}$$

Obs att moden existerar bara då $p, q > 1$.

Beta-fördelningen är lämplig som prior-täthet för en sannolikhet θ eftersom

$$\theta \sim \text{beta}(p, q), \quad x|\theta \sim \text{bin}(n, \theta)$$

implicerar

$$\theta|x \sim \text{beta}(p+x, q+n-x)$$

I vår tillämpning har vi prior-kunskapen representerad av

$$\theta \sim \text{beta}(1, 1) = U(0, 1)$$

Efter $n_1 = 7$ oberoende försök av vilka A inträffade i $x_1 = 4$ st, är den

$$\theta|n_1, x_1 \sim \text{beta}(5, 4)$$

Det är naturligt att skatta θ med posteriorns mod,

$$\hat{\theta} = \frac{4}{7} \approx 0.571$$

Ett s.k kreditibilitetsintervall med kreditibiliteten 0.95 är

$$0.245 \leq \theta \leq 0.843$$

ty m.a.p beta(5, 4)-fördelningen gäller

$$P(\theta < 0.245) = 0.025 \quad \text{och} \quad P(\theta > 0.843) = 0.025$$

Gränsema i kreditibilitetsintervallet kan t.ex räknas ut med Matlab-rutinen **betainv**.

Efter ytterligare $n_2 = 10$ oberoende försök av vilka A inträffade i $x_2 = 6$ st, är posteriorin

$$\theta|n_1, x_1, n_2, x_2 \sim \text{beta}(11, 8)$$

Nu är posterior-skattningen

$$\hat{\theta} = \frac{10}{17} \approx 0.588$$

och kreditibilitetsintervallet är

$$0.357 \leq \theta \leq 0.785$$

Kreditibiliteten är 0.95 även i detta intervall.

Obs att det i detta fall då priorn är beta(1, 1) = U(0, 1) inte finns någon prior-skattning. Annars hade moden i priorn varit det naturliga valet.

Har man inte möjlighet att bestämma gränserna i kreditibilitetsintervallet, kan man istället ange θ 's standardavvikelse. Efter den första mätserien är standardavvikelsen 0.157 och efter den andra har den minskat till 0.110.

När man har hyfsat många mätningar och $\hat{\theta}$ ej ligger för nära 0 eller 1, kommer p.g.a c.g.s gränsema i ett 95% kreditibilitetsintervall att bli ungefär

$$\mu \pm 2\sigma$$

där $\mu = E[\theta] \approx \hat{\theta}$ och $\sigma^2 = \text{Var}[\theta]$.

En ingenjör ska medelst provtagning avgöra hur stor sannolikheten θ är att ett visst gränsvärde överstigs. Hon gör därför 8 mätningar (som vi ska anta är oberoende upprepningar av försöket att bestämma en viss halt av någon förorening). I 3 av mätningarna överstegs gränsvärdet. Föreslå à priori-täthet för θ .

Anta att ingenjören tidigare gjort 4 mätningar och att gränsvärdet överstegs i en av dessa. Föreslå en à priori-täthet för θ som tar hänsyn till detta.

Anta istället att ingenjören erfarenhetsmässigt vet att $\theta \approx 0.25$ och att hon vill använda sig av denna erfarenhet i skattandet av θ . Hon bestämmer sig för att värdera sin erfarenhet lika mycket som den information hon erhåller från de 8 nya mätningarna. Föreslå ny à priori-täthet.

Istället för att jämföra värdet av nya data med redan erhållna skulle hon kunna tänka sig att ansätta $P(\theta > 0.10) = 0.9$. Föreslå ny à priori-täthet.

Etc.

Observera att man behöver ha två villkor för att härleda lämpliga värden på parametrarna p, q i à priori-fördelningen.