הסקה סטטיסטית שיעור 6

הרצאה [מצגת שיעור 6](),

1 עדכון בייסיאני: קודמים רציפים

הגישה הבייסיאנית מניחה prior לגבי ההשערות, ואז בהינתן דאטא מעדכנת את המשקולות של כל השערה באמצעות חוק בייס.

חישוב האינטגרל הוא הגורם המגביל העיקרי. שמנו לב שאם מסתכלים על צורה מסויימת של הprior אז הposterior הוא גם מאותה משפחה. לדוגמה ראינו התפלגות Beta: θ יכול לקבל ערכים בין 0 ל-1.

$$f(\theta)=\frac{(a+b-1)!}{(a-1)!\,(b-1)!}\theta^{a-1}(1-\theta)^{b-1}$$

וראינו שהמקדם הוא גורם מנרמל, כלומר אם יש לנו PDF:

$$f(\theta)=c\theta^{a-1}(1-\theta)^{b-1}$$

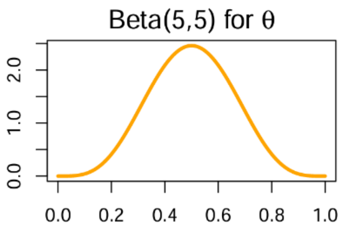
אז $\theta\sim Beta(a,b)$, ומתקיים:

$$c=\frac{(a+b-1)!}{(a-1)!\,(b-1)!}$$

אם $a=1,b=1$ אז ההתפלגות היא $f(\theta)=1$.

אם $a=2,b=2$ אז זה מבנה של פרבולה.

אם $a>b$ זה נוטה שמאלה, אם $a<b$ זה נוטה ימינה.



תרגיל כיתה 1

נניח שאנחנו בודקים תרופה עם הסתברות לא ידועה θ להצלחה. אנחנו לא יודעים את θ אבל אנחנו מנחשים שהוא באיזור חצי. נראה את האינטואיציה הזו באמצעות prior של $Beta(5,5)$ על θ. כלומר:

נניח שבדקנו 10 מטופלים וקיבלנו 6 הצלחות. נמצא את ההתפלגות הposterior של תטא:

הPDF הprior היא:

$$f(\theta)=\frac{9!}{4!\,4!}\theta^4(1-\theta)^4$$

hypoth.	prior	likelihood	Bayes numer.	posterior
θ	$c_1\theta^4(1-\theta)^4\,d\theta$	$\binom{10}{6}\theta^6(1-\theta)^4$	$c_3\theta^{10}(1-\theta)^8\,d\theta$	beta$(11,9)$

ה-likelihood זה פשוט ההסתברות לקבל 6 הצלחות מתוך 10.

הposterior, לפי הגדרה:

$$f(\theta|D)d\theta=\frac{p(D|\theta)f(\theta)\,d\theta}{p(D)}=\frac{p(D|\theta)f(\theta)}{\int_a^b p(D|\theta)f(\theta)\,d\theta}=\frac{c_3\theta^{10}(1-\theta)^8}{\int_0^1 c_3\theta^{10}(1-\theta)^8d\theta}$$

זה עדיין התפלגות בטא, עם פרמטרים שונים.

ההסתברות להצלחה בניסיון הבא:

$$P(S|D)=\int_0^1 f(S|\theta)f(\theta|D)d\theta=\int_0^1\theta\frac{19!}{10!\,8!}\theta^{10}(1-\theta)^8\,d\theta=\int_0^1\frac{19!}{10!\,8!}\theta^{11}(1-\theta)^8\,d\theta$$

נשתמש במידע שלנו לגבי הPDF כדי להעריך את האינטגרל בלי לחשב אותו ישירות:

נשווה אותו לPDF של $Beta(12,9)$. נוציא את המקדם, נכפיל ונחלק. אנחנו יודעים:

$$\frac{(a+b-1)!}{(a-1)!\,(b-1)!}\int_0^1\theta^{a-1}(1-\theta)^{b-1}\,d\theta=1$$

אז:

$$= \frac{19!}{10!\,8!}\frac{11!\,8!}{20!}\underbrace{\frac{20!}{11!\,8!}\int_0^1 \theta^{11}(1-\theta)^8\,d\theta}_{f(\theta)\sim Beta(12,9),\ \ =1} = \frac{11}{20}$$

התפלגות צמודה

היה לנו: prior $f(\theta)d\theta$ התפלגות בטא, likelihood $p(x|\theta)$ התפלגות בינומית, posterior $f(\theta|x)d\theta$ התפלגות בטא.

להתפלגות בטא קוראים **התפלגות צמודה** עבור הlikelihood הבינומית. ה-prior בטא הופכת ל-posterior בטא וקל לעשות עדכונים חוזרים.

בעצם ראינו, שאם יש prior $Beta(a,b)$, והטלנו $i+j$ פעמים וקיבלנו i עץ, אז ה posterior הוא $Beta(a+i,b+j)$.

תזכורת: הסתברויות חיזוי

Prior predictive probability: $p(x_1) = \int p(x_1|\theta)f(\theta)d\theta$.

Posterior predictive probability: $p(x_2|x_1) = \int p(x_2|\theta)f(\theta|x_1)d\theta$.

קודמים רציפים, דאטא רציפה:

$$f(x) = \int p(x|\theta)f(\theta)d\theta$$

hypoth.	prior	likelihood	Bayes numerator	posterior $f(\theta\|x)\,d\theta$
θ	$f(\theta)\,d\theta$	$f(x\,\|\,\theta)$	$f(x\,\|\,\theta)f(\theta)\,d\theta$	$\frac{f(x\,\|\,\theta)f(\theta)\,d\theta}{f(x)}$
total	1		$f(x)$	1

תרגיל כיתה: קודם נורמלי, דאטא נורמלית

ל- $N(\mu,\sigma^2)$ יש את הצפיפות:

$$f(y) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right)$$

נשים לב: המקדם הוא גורם מנרמל, אז אם יש את הPDF:

$$f(y) = c\exp\left(-\frac{(y-\mu)^2}{2\sigma^2}\right)$$

אז $y\sim N(\mu,\sigma^2)$, וגם:

$$c = \frac{1}{\sigma\sqrt{2\pi}}$$

נניח שהדאטא שלנו מתפלגת $N(\theta,4)$. כלומר $f(x|\theta) = pdf\ of\ N(\theta,4)$.

נניח שה prior שלנו על θ היא $N(3,1)$.

נניח שקיבלנו את הדאטא $x_1 = 5$. נמצא את ה posterior PDF של θ.

יש לנו: *prior*: $\theta\sim N(3,1)$. $f(\theta) = c_1\exp(-(0-3)^2/2)$.

Likelihood: $x\sim N(\theta,4)$. $f(x|\theta) = c_2\exp(-(x-\theta)^2/8)$.

עבור $x=5$ ה *likelihood* היא $c_2\exp(-(5-\theta)^2/8)$.

hypoth.	prior	likelihood	Bayes numer.
θ	$c_1 e^{-(\theta-3)^2/2}\,d\theta$	$c_2 e^{-(5-\theta)^2/8}\,dx$	$c_3 e^{-(\theta-3)^2/2}e^{-(5-\theta)^2/8}\,d\theta\,dx$

נפשט את הביטוי:

$$c_3 \exp\left(-\frac{(\theta-3)^2}{2}\right)\exp\left(-\frac{(5-\theta)^2}{8}\right)d\theta dx = c_3 \exp\left(-\frac{5}{8}\left(\theta^2-\frac{34}{5}\theta+61\right)\right)$$

השלמה לריבוע:

$$\exp\left(-\frac{5}{8}\left(\theta^2-\frac{34}{5}\theta+61\right)\right)$$

נתחיל עם משהו שנותן לנו את θ^2 ואת $-\frac{34}{5}$:

$$\left(\theta-\frac{17}{5}\right)^2 = \theta^2-\frac{34}{5}\theta+\left(\frac{17}{5}\right)^2$$

עכשיו צריך להוריד $\left(\frac{17}{5}\right)^2$, ולהוסיף 61:

$$\left(\theta-\frac{17}{5}\right)^2-\left(\frac{17}{5}\right)^2+61$$

בסה"כ:

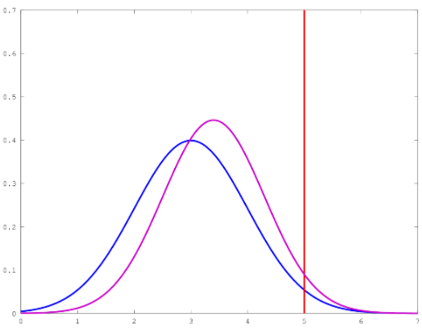
$$\exp\left(-\frac{5}{8}\left(\left(\theta-\frac{17}{5}\right)^2-\left(\frac{17}{5}\right)^2+61\right)\right)$$

כלומר:

$$c_3 \exp\left(-\frac{5}{8}\left(\left(\theta-\frac{17}{5}\right)^2-\left(\frac{17}{5}\right)^2+61\right)\right) = c_3 \exp\left(-\frac{5}{8}\left(61-\left(\frac{17}{5}\right)^2\right)\right)\exp\left(-\frac{5}{8}\left(\theta-\frac{17}{5}\right)^2\right)$$
$$= \mathrm{c}_4 \exp\left(-\frac{5}{8}\left(\theta-\frac{17}{5}\right)^2\right) = c_4 \exp\left(-\frac{\left(\theta-\frac{17}{5}\right)^2}{2\cdot\frac{4}{5}}\right)$$

מה שבעצם רצינו זה להגיע למצב שכל הקבועים מבודדים בתור מקדם, ויש לנו $(\theta-z)^2$ במעריך של e.

וזה אומר לנו שה posterior היא:



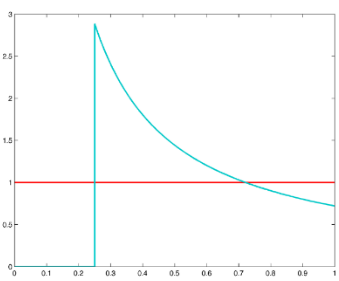
$$\theta \sim N\left(\frac{17}{5},\frac{4}{5}\right)$$

Prior: כחול, posterior: סגול, דאטא: אדום.

$$prior \sim N(3,1), \qquad likelihood \sim N(\theta,4), \qquad posterior \sim N\left(\frac{17}{5},\frac{4}{5}\right)$$

נראה בהמשך טכניקות קלות יותר לעדכונים כאלה.

תרגיל כיתה: רומיאו ויוליה



רומיאו מאחר תמיד בהתפלגות אחידה $U(0,\theta)$, עם פרמטר לא ידוע θ. כלומר הוא מאחר בין 0 ל-θ שעות.

יוליה יודעת ש $1 \geq \theta$, ומניחה flat prior (קודמת אחידה) עבור θ, על $[0,1]$.

בדייט הראשון רומיאו מאחר ב15 דקות. נעדכן את הprior של θ.

נחלק את ההשערות לשניים: $\theta < 0.25, \theta \geq 0.25$.

hyp.	prior $f(\theta)$	likelihood $f(x_1\|\theta)$	Bayes numerator	posterior $f(\theta\|x_1)$
$\theta < 0.25$	$d\theta$	0	0	0
$\theta \geq 0.25$	$d\theta$	$\frac{1}{\theta}$	$\frac{d\theta}{\theta}$	$\frac{c}{\theta}\, d\theta$
Tot.	1		T	1

וצריך שהקבוע המנרמל c יגרום לכל ה$posterior$ להיות 1, כלומר:

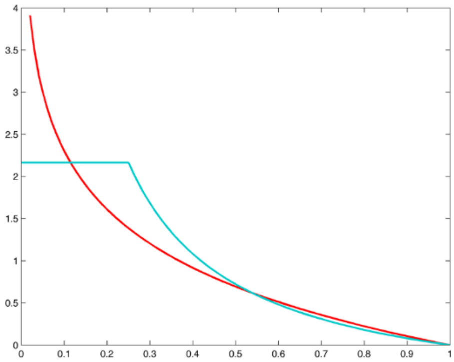
$$c\int_{0.25}^{1} d\theta = 1 \quad \Rightarrow \quad c = \frac{1}{\ln 4}$$

אז ה$prior\ predictive\ PDF$ של x_1 הוא:

$$f(x_1) = \int_0^1 f(x_1|\theta)f(\theta)d\theta = \int_{x_1}^1 \frac{1}{\theta}d\theta = -\ln(x_1)$$

חיזוי אוחר $posterior\ prediction$:

$$f(x_2|\theta) = \begin{cases} \frac{1}{\theta}, & if\ \theta \geq x_2 \\ 0, & if\ \theta < x_2 \end{cases}$$



הposterior predictive PDF $f(x_2|x_1)$ הוא: $\int_0^1 f(x_2|\theta)f(\theta|x_1)d\theta$

שזה 0 אלא אם כן $\theta > x_2, \theta > 0.25$. יש שני מקרים:

אם $x_2 < 0.25$: $f(x_2|x_1) = \int_{0.25}^1 \frac{c}{\theta^2}d\theta = 3c = \frac{3}{ln4}$

אם $x_2 \geq 0.25$: $f(x_2|x_1) = \int_{x_2}^1 \frac{c}{\theta^2}d\theta = \frac{\frac{1}{x_2}-1}{ln4}$

תרגול

נתון לנו: $\mu \sim N(\mu, \sigma^2)$. נחשב את ה posterior: $P(\mu|x_1, \ldots x_n) \propto \{(x_1, \ldots x_n) \cdot P(\mu)\}$

נתחיל עם נקודה אחת: $P(X|\mu) \sim N(\mu, \sigma^2)$

$$P(\mu|X) \propto \mathcal{L}(X) \cdot P(\mu) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{(X-\mu)^2}{2\sigma^2}\right) \cdot \frac{1}{\sigma_0\sqrt{2\pi}}\exp\left(-\frac{(\mu-\mu_0)^2}{2\sigma_0^2}\right)$$

$$\propto \exp\left(-\frac{1}{2}\left[\frac{1}{\sigma^2}(X-\mu)^2 + \frac{1}{\sigma_0^2}(\mu-\mu_0)^2\right]\right) =$$

$$= \exp\left(-\frac{1}{2}\left[\frac{1}{\sigma^2}(X^2 - 2\mu X + \mu^2) + \frac{1}{\sigma_0^2}(\mu^2 - 2\mu\mu_0 + \mu_0^2)\right]\right) =$$

$$= \exp\left(-\frac{1}{2}\left[\left(\frac{1}{\sigma^2} + \frac{1}{\sigma_0^2}\right)\mu^2 \cdot 2\left(\frac{X}{\sigma^2} + \frac{\mu_0}{\sigma_0^2}\right)\mu + \frac{X^2}{\sigma^2} + \frac{\mu_0^2}{\sigma_0^2}\right]\right) =$$

אנחנו רוצים שזה יהיה בצורה של:

$$\exp\left(-\frac{1}{2}[\_\_(\mu - \_\_)^2]\right)$$

השלמה לריבוע, כלומר:

$$= \exp\left(-\frac{1}{2}\left[a\left(\mu - \frac{b}{2a}\right)^2 + c - \frac{b^2}{4a}\right]\right) \propto \exp\left(-\frac{a}{2}\left(\mu - \frac{b}{2a}\right)^2\right) \sim N\left(\frac{b}{2a}, \frac{1}{a}\right)$$

כלומר גם ה $posterior$ מתפלג נורמלית. (התפלגות צמודה).

$$\mathcal{L}(x_1, \ldots, x_n) = \frac{1}{\sigma^n(2\pi)^{n/2}}\exp\left(\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2\right)$$

אנחנו יודעים:

$$\sum_{i=1}^{n}(x_i-\mu)^2=\left(\sum_{i=1}^{n}(x_i-\bar{X})^2\right)-n(\bar{X}-\mu)^2$$

והחלק בסוגריים הוא קבוע (לא תלוי ב- μ) אז:

$$\propto \exp\left(\frac{1}{2\sigma/n}(\bar{X}-\mu)^2\right)\sim N\left(\mu,\frac{\sigma^2}{n}\right)$$

וה *posterior*:

$$N\left(\frac{\frac{\bar{x}}{\sigma^2/n}+\frac{\mu_0}{\sigma_0^2}}{\frac{1}{\sigma^2/n}+\frac{1}{\sigma_0^2}},\qquad \frac{1}{\frac{1}{\sigma^2 n}+\frac{1}{\sigma_0^2}}\right)$$