DATE / / SUBJECT:

سوال ۱:

(A, B

$x_i \sim \text{Binomial}(N_i, \theta)$

$\theta \sim \text{Beta}(\alpha, \beta)$

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

$$\propto \binom{N}{x_i} \theta^S (1-\theta)^{N-S} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha-1} (1-\theta)^{\beta-1}$$

$$\propto \theta^S (1-\theta)^{N-S} \theta^{\alpha-1} (1-\theta)^{\beta-1}$$

$S = \sum_{i=1}^{n} x_i$

$N = \sum_{i=1}^{n} N_i$

$$\alpha^* = \alpha + \sum_{i=1}^{n} x_i$$

$$\beta^* = \beta + \sum_{i=1}^{n} N_i - \sum_{i=1}^{n} x_i$$

posterior $\propto$ prior $\times$ likelihood

Binomial likelihood $\propto p^x (1-p)^{n-x}$

Beta prior $\propto p^{\alpha-1} (1-p)^{\beta-1}$

Beta posterior $\propto p^x (1-p)^{n-x} p^{\alpha-1} (1-p)^{\beta-1}$

DATE / / SUBJECT:

Beta posterior $\propto p^{x+\alpha-1}(1-p)^{n-x+\beta-1}$

DATE / / SUBJECT:

سوال ۲ :

H1: good die on the right

با دو فرض شروع می کنیم

H2: good die on the left

ابتدا هر دو ۵۰٪ احتمال دارند.

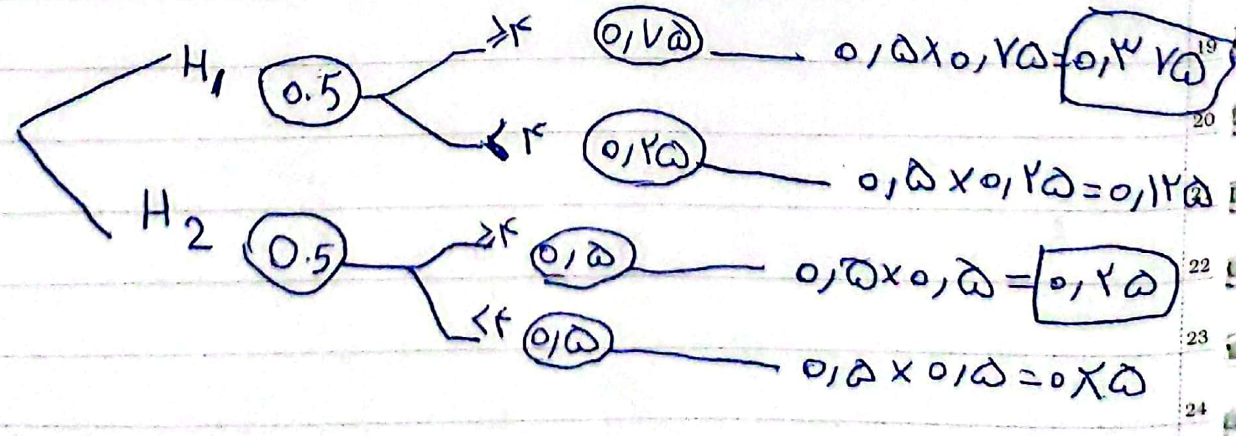
prior probability

سپس در مرحله بعد انتخاب می کنیم (دوره ای)

★ درصورتی که H_1 درست باشد (تاس خوب در دست راست)، احتمال این که عددی ≥۴ باید ۷۵٪ می شود و عدد کمتر از ۴ ۲۵٪ می شود.

★ در غیر اینصورت، اگر تاس بد در دست راست باشد، احتمال عدد بزرگتر از ≥۴ (مساوی) باید ۵۰٪ و کمتر از ۴ هم ۵۰٪ است.

اگر احتمالات را ضرب کنیم به این شکل می شود:



ادامه سوال ۲:

$$P(H_1: \text{good die on the right} \mid \text{rolled} \geq 4 \text{ with right})$$

$$= \frac{P(\text{good in Right} \;\&\; \text{rolled} \geq 4 \text{ right})}{P(\text{rolled} \geq 4 \text{ right})}$$

$$= \frac{۰٫۳۷۵}{۰٫۳۷۵ + ۰٫۲۵} = ۰٫۶$$

سوال ۲ قسمت c: در روش bayesian با احتمالات و متغیر تصادفی سروکار داریم و در روش frequentist پارامتر ها را ثابت می‌گیریم و در آن به مشاهدات و شواهد آن ها استناد می شود.

ولی در بیزین "مقدار اولیه" هم مهم است و در نتیجه گیری مورد استفاده قرار می گیرد.

سوال ۳:

$$f_X(x) = \frac{1}{\sqrt{2\pi}\,\sigma_X} e^{-\frac{x^2}{2\sigma_X^2}}$$

$$Y|X=x \sim N(x, \sigma_W^2) \rightarrow f_{Y|X}(y|x) = \frac{1}{\sqrt{2\pi}\,\sigma_W} e^{-\frac{(y-x)^2}{2\sigma_W^2}}$$

ML

$$f_{Y|X}(y|x) = \frac{1}{\sqrt{2\pi}\,\sigma_W} e^{-\frac{(y-x)^2}{2\sigma_W^2}}$$ باید این تابع را بیشینه کنیم

پس $(y-x)^2$ را کمینه می‌کنیم (چون توان منفی است)

$$\rightarrow \hat{x}_{ML} = y.$$

MAP

$$f_{Y|X}(y|x) f_X(x) = c \exp\left\{-\left[\frac{(y-x)^2}{2\sigma_W^2} + \frac{x^2}{2\sigma_X^2}\right]\right\}$$

$$\frac{(y-x)^2}{2\sigma_W^2} + \frac{x^2}{2\sigma_X^2} \rightarrow \hat{x}_{MAP} = \frac{\sigma_X^2}{\sigma_X^2 + \sigma_W^2} y$$

سوال ۴:

$$\begin{cases} Y = 2X + W \\ W \sim N(0, \sigma^2) \\ X = 1 \quad p \\ X = -1 \quad 1-p \end{cases}$$

$H_0 : X = 1,$

$H_1 : X = -1$

$$H_0 \to \quad Y = 2 + W$$

$$\to Y \mid H_0 \sim N(2, \sigma^2)$$

$$\to \quad f_Y(y \mid H_0) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y-2)^2}{2\sigma^2}}$$

$$H_1 \to \quad Y = -2 + W$$

$$\to Y \mid H_1 \sim N(-2, \sigma^2)$$

$$\to f_Y(y \mid H_1) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y+2)^2}{2\sigma^2}}$$

پس فقط درصورتی H_0 را انتخاب می‌کنیم که

$$\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y-2)^2}{2\sigma^2}} P(H_0)$$

$$\geq \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(y+2)^2}{2\sigma^2}} P(H_1)$$

چون $P(H_0) = P$ و $P(H_1) = 1-p$

$$\to e^{\left(\frac{4y}{\sigma^2}\right)} \geq \frac{1-p}{p}$$

که معادل: $y \geq \frac{\sigma^2}{4} \ln\left(\frac{1-p}{p}\right)$

سوال ۵:

از الگوریتم Metropolis-Hastings برای تقریب زدن توزیع ها بوسیله زنجیر مارکوف مونت کارلو استفاده می شود. این الگوریتم روش بسیار قدرتمندی است. این الگوریتم فقط به یک عبارت نیاز دارد که متناسب با چگالی نمونه است که معمولا صورت چگالی posterior است.

فرض کنید توزیعی متناسب با تابع f داریم. شبه کد را اینطور می نویسیم:

۱- θ_0 مقداردهی اولیه شود

۲- به ازای i از ۰ تا N:

۱- از $q(x|\theta_{i-1})$ ، θ_i^* را نمونه برداری کن

۲- مقدار θ_i را برابر با θ_i^* با این احتمال قرار بده:

$$\alpha = \min\left(\frac{f(\theta_i^*)\, q(\theta_{i-1}|\theta_i^*)}{f(\theta_{i-1})\, q(\theta_i^*|\theta_{i-1})}, 1\right)$$

در غیر اینصورت $\theta_i=\theta_{i-1}$

که این مقدار α می تواند به مقدار $\alpha = \min\left(\frac{f(\theta_i^*)}{f(\theta_{i-1})}, 1\right)$ ساده تر شود در صورتی که از توزیع نرمال q استفاده شود.

سوال ۶

A) علت این که این مسئله متناقض به نظر می آید این است که دقت یک تست (sensitivity) نشان دهنده‌ی قابلیت پیش بینی کردن آن تست (predictablity) نیست.

$$\text{accuracy} \not\Rightarrow \text{preditability}$$

برای مثال در نمونه ۱۰۰۰ نفره که احتمال داشتن سرطان ۱۰ درصد است، اگر تستی داشته باشیم که دارای دقت ۹۰٪ است لزوما نمی توانیم بگوییم که فردی که تست + دارد با احتمال ۹۰٪ واقعا سرطان دارد. (با این که شهودی این طور به نظر می آید)

روش سنتی احتمال ها را در نظر می گیرد ولی روش جدیدتر از odds استفاده می کند که برای انسان قابل فهم تر است.

(که در قسمت های بعدی بیشتر توضیح داده شده)

B) این ویدئو ابتدا ورژن دیگر فرمول بیز را بیان می‌کند:

فرمول اصلی :
$$P(H|E) = \frac{P(H)\,P(E|H)}{P(E)}$$

ورژن جدید :
$$\underbrace{O(H|E)}_{\text{odds}} = O(H)\,\underbrace{\frac{P(E|H)}{P(E|\neg H)}}_{\text{bayes factor}}$$

فرض کنید ۱۰۰۰ نفر را درنظر می‌گیریم که ۱٪ از آن‌ها سرطان دارند

یک تست موردنظر می‌تواند با احتمال ۹۰٪ تشخیص دهد که شخص بیمار نتیجه + دارد.

$P(+|\text{cancer}) = \%90$ ← sensitivity

این تست همچنین برای افراد سالم اینطور تشخیص میدهد:

$P(+|\neg\text{cancer}) = \%9$

$\to P(-|\neg\text{cancer}) = \%91$ ← specifity

با درنظر گرفتن مقدار sensitivity و specificity تست

که هردو بالای ۹۰٪ اند می بینیم که تست ما دقت بالایی دارد. اما

خود احتمال ابتلا به سرطان (prevalence) خیلی کم است (۱٪)

سوال این است که اگر یک نفر تست + داشته باشد احتمال این که سرطان دارد چقدر است؟ چون دقت تست بالا است در نگاه اول فکر می‌کنیم که از هر ۱۰ نفر ⊕ ۹ نفرشان سرطان دارند ولی در واقع این طور نیست و باید این که احتمال سرطان کم است را هم در نظر بگیریم.

ابتدا bayes factor معرفی می‌شود که برابر است با

$$\frac{P(+|\text{cancer})}{P(+|\text{No cancer})} = \frac{90\%}{9\%} = 10$$

این مقدار هرچه بیشتر باشد دقت تست بالاتر است.

در ادامه روش فرمول جدید که با odds است معرفی می‌شود. در استفاده از odds به جای نشان دادن احتمال، از نسبت تعداد افراد مورد نظر استفاده می‌کنیم یعنی مثلا

۱۰٪ می‌شود 10:90 (۱۰ به ۹۰)

۲۰٪ می‌شود ۲۰:۸۰ (۲۰ به ۸۰)

در متد odds ابتدا (۱) مقدار prior (احتمال سرطان داشتن کل افراد) را به شکل odds می‌نویسیم:

مثلا اگر prior ۱٪ باشد ← 1:99

DATE / / SUBJECT:

۲) سپس bayes factor را محاسبه می کنیم :

$$\frac{P(+|cancer)}{P(+|No\,cancer)} = \frac{90\%}{9\%} = 10$$

۳) ضرب می کنیم $(1:99)\times 10 = 10:99 \rightarrow \frac{10}{109} \simeq \frac{1}{11}$

پس یعنی احتمال این که شخصی که تست + دارد واقعا سرطان

داشته باشد $\frac{1}{11}$ است. $p(cancer|+) = \frac{1}{11}$

$$\text{bayes factor} = \frac{\text{sensitivity}}{1-\text{specifity}}$$

بیشتر در قسمت B CC توضیح داده شد

در روش قبلی با احتمال ها سروکار داریم ولی با دانستن bayes factor

به سادگی می توان میزان $PPV = P(cancer|+)$ را با دانستن

prior و bayes factor محاسبه کرد و از حجم محاسبات و پیچیدگی کاسته می شود

(D

prior: %۸ → ۸ : ۹۲

Sensitivity = %۸۰

~~P(not infected | +)~~ = 1 : 9 → %۱۰

P(+ | not infected) =

question: P(infected | +) = ?

جواب:

bayes factor = $\frac{\%۸۰}{\%۱۰}$ = ۸

→ prior × bayes Factor = (۸ : ۹۲) × ۸ = (۶۴ : ۹۲)

→ $\frac{۶۴}{۹۲+۶۴}$ = %۴۱