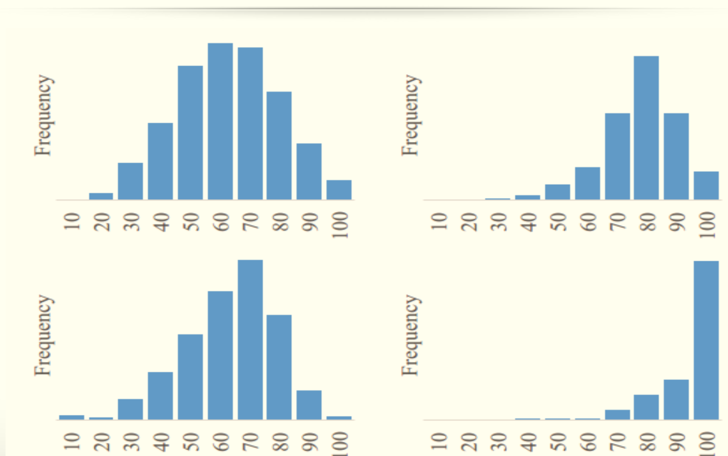
آمار و احتمال مهندسی

توزیع بتا

(Ross 5.6, 6.5)

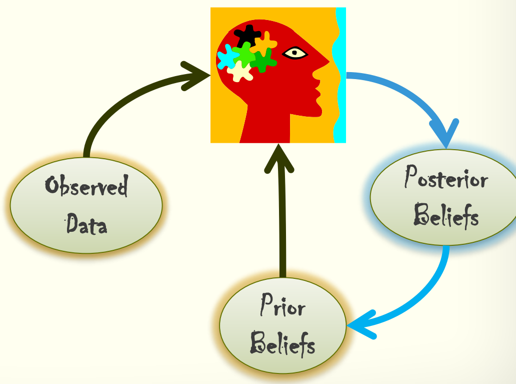
توزیع نمرات تمرین‌ها



تعبیر بیزی احتمال

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}$$

- $P(B)$: میزان باور ابتدایی ما نسبت به B (احتمال پیشین)
- $P(B|A)$: میزان باور ما نسبت به B پس از مشاهده A (احتمال پسین)
- $P(A|B)/P(A)$: میزان حمایت A از B



قضیه بیز

- قضیه بیز برای توابع جرمی احتمال:

$$P_{Y|X}(y|x) = \frac{P_{XY}(x,y)}{P_X(x)} = \frac{P_{X|Y}(x|y)P_Y(y)}{\sum_{y=-\infty}^{\infty} P_{X|Y}(x|y)P_Y(y)}$$

- قضیه بیز برای توابع چگالی احتمال:

$$f_{X|Y}(x|y) = \frac{f_{Y|X}(y|x)f_X(x)}{f_Y(y)} = \frac{f_{Y|X}(y|x)f_X(x)}{\int_{-\infty}^{+\infty} f_{Y|X}(y|x)f_X(x)dx}$$

ترکیب شرطی متغیرهای پیوسته و گسسته

- فرض کنید X یک متغیر تصادفی پیوسته و N یک متغیر تصادفی گسسته باشد.
- تابع چگالی احتمال شرطی X به شرط داشتن N به صورت زیر تعریف می‌شود:

$$f_{X|N}(x|n) = \frac{P_{N|X}(n|x) f_X(x)}{P_N(n)}$$

- تابع جرمی احتمال شرطی N به شرط داشتن X به صورت زیر تعریف می‌شود:

$$P_{N|X}(n|x) = \frac{f_{X|N}(x|n) P_N(n)}{f_X(x)}$$

- اگر X و N مستقل از یکدیگر باشند، آنگاه:

$$f_{X|N}(x|n) = f_X(x), \qquad P_{N|X}(n|x) = P_N(n)$$

تابع گاما

- تابع گاما به صورت زیر تعریف می‌شود:

$$\Gamma(r) = \int_0^\infty y^{r-1} e^{-y} \mathrm{dy}$$

- در واقع تابع $\Gamma(\cdot)$ فاکتوریل تعمیم‌یافته است. چون $\Gamma(r) = (\mathrm{r}-1)\Gamma(r-1)$، و $\Gamma(1) = 1$، برای r صحیح و مثبت داریم: $\Gamma(r) = (\mathrm{r}-1)!$

- در حالت خاص $r = 1/2$:

$$\Gamma(1/2) = \sqrt{\pi}$$

متغیر تصادفی بتا

- متغیر تصادفی پیوسته X را متغیر بتا می‌گوییم و با $X \sim Beta(a,b)$ نمایش می‌دهیم، اگر تابع چگالی احتمال آن به شکل زیر باشد:

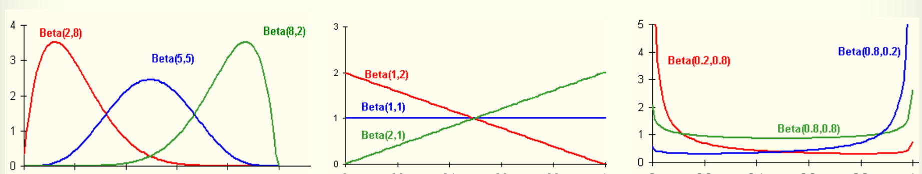
$$f_X(x) = \begin{cases} \dfrac{1}{\beta(a,b)} x^{a-1}(1-x)^{b-1} & 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

- که تابع بتا $\beta(a,b)$ به صورت زیر تعریف می‌شود:

$$\beta(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1}\,\mathrm{d}x = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}$$

$$\Gamma(n) = \int_0^\infty y^{n-1}e^{-y}\mathrm{d}y \quad \Rightarrow \quad \Gamma(n) = (n-1)!$$

توزیع بتا



- اگر $a = b = 1$ توزیع بتا تبدیل به توزیع یکنواخت روی $[0,1]$ می شود.
- اگر $a = b$، شکل تابع چگالی احتمال توزیع بتا متقارن می‌شود.
- اگر $X \sim Beta(a,b)$:

$$E[X] = \frac{a}{a+b}, \qquad Var(X) = \frac{ab}{(a+b)^2(a+b+1)}$$

میانگین توزیع بتا

- میانگین توزیع بتا به صورت زیر محاسبه می‌شود:

$$E[X] = \int_{-\infty}^{\infty} x\, f_X(x)dx = \int_0^1 x \frac{1}{\beta(a,b)} x^{a-1}(1-x)^{b-1}\ dx$$

$$= \frac{1}{\beta(a,b)} \int_0^1 x^{(a+1)-1}(1-x)^{b-1}\ dx$$

$$= \frac{1}{\beta(a,b)} \times \beta(a+1,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \times \frac{\Gamma(a+1)\Gamma(b)}{\Gamma(a+b+1)}$$

$$= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \times \frac{a\Gamma(a)\Gamma(b)}{(a+b)\Gamma(a+b)} = \frac{a}{a+b}$$

واریانس توزیع بتا

- واریانس توزیع بتا به صورت زیر محاسبه می‌شود:

$$E[X^2] = \int_{-\infty}^{\infty} x^2\, f_X(x)dx = \int_0^1 x^2 \frac{1}{\beta(a,b)} x^{a-1}(1-x)^{b-1}\ dx$$

$$= \frac{1}{\beta(a,b)} \int_0^1 x^{(a+2)-1}(1-x)^{b-1}\ dx$$

$$= \frac{1}{\beta(a,b)} \times \beta(a+2,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \times \frac{\Gamma(a+2)\Gamma(b)}{\Gamma(a+b+2)}$$

$$= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \times \frac{(a+1)a\Gamma(a)\Gamma(b)}{(a+b+1)(a+b)\Gamma(a+b)} = \frac{a(a+1)}{(a+b)(a+b+1)}$$

واریانس توزیع بتا

- واریانس توزیع بتا به صورت زیر محاسبه می‌شود:

$$Var(X) = E[X^2] - (E[X])^2$$
$$= \frac{a(a+1)}{(a+b)(a+b+1)} - \left(\frac{a}{a+b}\right)^2$$
$$= \frac{a}{a+b}\left(\frac{a+1}{a+b+1} - \frac{a}{a+b}\right)$$
$$= \frac{a}{a+b} \times \frac{(a+1)(a+b) - a(a+b+1)}{(a+b)(a+b+1)}$$
$$= \frac{a}{a+b} \times \frac{b}{(a+b)(a+b+1)} = \frac{ab}{(a+b)^2(a+b+1)}$$

مد توزیع بتا

- مد یک توزیع پیوسته، برابر با نقطه ماکزیمم تابع چگالی احتمال آن است:

$$f_X(x) = \frac{1}{\beta(a,b)} x^{a-1}(1-x)^{b-1}$$

$$\frac{df_X(x)}{dx} = \frac{1}{\beta(a,b)}\left((a-1)x^{a-2}(1-x)^{b-1} - (b-1)x^{a-1}(1-x)^{b-2}\right) = 0$$

$$\frac{1}{\beta(a,b)} x^{a-2}(1-x)^{b-2}\left((a-1)(1-x) - (b-1)x\right) = 0$$

$$(a-1) - (a+b-2)x = 0 \Rightarrow x = \frac{a-1}{a+b-2}$$

پرتاب سکه با احتمال ناشناخته

- کاربرد اصلی توزیع بتا در نمایش توزیع باور ما درباره یک احتمال است.
- سکه‌ای را $(n+m)$ بار پرتاب می‌کنیم و n بار شیر می‌آید. احتمال شیر آمدن سکه یا X را نمی‌دانیم:

تعبیر بیزی احتمال
(Bayesian)

$$f_{X|N}(x|n) = \frac{P(N=n|X=x)f_X(x)}{P(N=n)}$$

X یک متغیر تصادفی است.

تعبیر بسامدی احتمال
(Frequentist)

$$X = \lim_{n+m\to\infty} \frac{n}{n+m} \approx \frac{n}{n+m}$$

X یک مقدار حقیقی است.

پرتاب سکه با احتمال ناشناخته

- اگر احتمال شیر آمدن سکه را X بنامیم، در ابتدا باور ما نسبت به این احتمال با توجه به اصل ناکافی بودن دلیل یکنواخت است:

$$X \sim U(0,1)$$

- متغیر تصادفی N را برابر تعداد شیرها در $n+m$ پرتاب تعریف می‌کنیم.
- اگر احتمال شیر آمدن را $X=x$ در نظر بگیریم، $(N|X)$ دارای توزیع دوجمله‌ای خواهد بود:

$$N|X \sim \mathrm{Bin}(n+m, x)$$

توزیع احتمال پیشین (prior):
$X \sim U(0,1)$

$$f_{X|N}(x|n) = \frac{P(N=n|X=x)f_X(x)}{P(N=n)}$$

توزیع احتمال پسین (posterior): ؟

محاسبه توزیع احتمال پسین

$$f_{X|N}(x|n) = \frac{P(N=n|X=x)f_X(x)}{P(N=n)}$$

$$= \frac{\binom{n+m}{n} x^n(1-x)^m \times 1}{P(N=n)} \quad : \quad 0<x<1$$

$$= \frac{\binom{n+m}{n}}{P(N=n)} x^n(1-x)^m \quad : \quad 0<x<1$$

$$= \frac{1}{c}\, x^n(1-x)^m \quad : \quad 0<x<1$$

محاسبه توزیع احتمال پسین

- از آنجا که $f_{X|N}$ یک توزیع احتمال است، باید داشته باشیم:

$$\int_{-\infty}^{\infty} f_{X|N}(x|n)\, dx = 1$$

$$\Rightarrow \int_0^1 \frac{1}{c} x^n(1-x)^m dx = 1$$

$$\Rightarrow \frac{1}{c}\int_0^1 x^n(1-x)^m\, dx = 1$$

$$\Rightarrow c = \int_0^1 x^n(1-x)^m\, dx = \beta(n+1, m+1)$$

محاسبه توزیع احتمال پسین

- از آنجا که $f_{X|N}$ یک توزیع احتمال است، باید داشته باشیم:

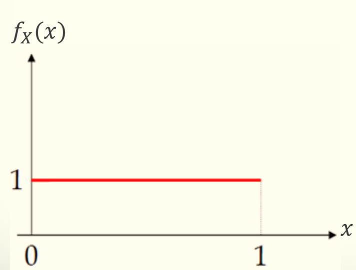
$$f_{X|N}(x|n) = \frac{1}{\beta(n+1, m+1)} x^n (1-x)^m \quad : \quad 0 < x < 1$$

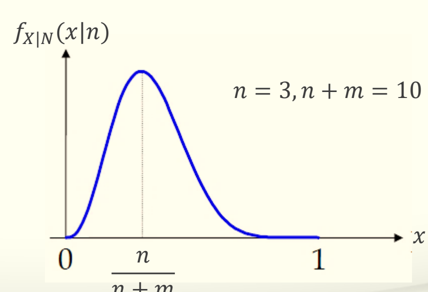
- بنابراین توزیع پسین متغیر تصادفی X پس از مشاهده $N = n$ یک توزیع بتا با پارامترهای $n+1$ و $m+1$ است:

$$X|(N = n) \sim Beta(n+1, m+1)$$

تخمین بیزی (Bayesian Estimation)

- یعنی X که توزیع یکنواخت داشت با مشاهده مقدار $N = n$ (تعداد شیرهای آمده در $n+m$ بار پرتاب سکه) توزیع بتا پیدا می‌کند.
- ماکزیمم توزیع بتا در $\frac{a-1}{a+b-2}$ یعنی $\frac{n}{n+m}$ اتفاق می‌افتد.
    - هرچه $n+m$ بزرگتر باشد، نمودار تیزتر می‌شود.

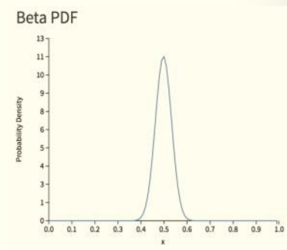




متغیر تصادفی بتا برای تخمین احتمال

۱۹۹ پرتاب و ۹۹ شیر

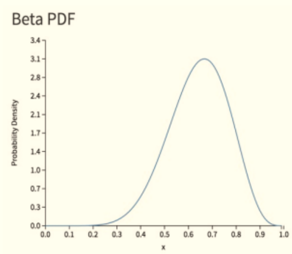
$f_{X\,|(100 \text{ heads},\ 101 \text{ tails})}$



$X \sim \text{Beta}(101{,}102)$

۱۲ پرتاب و ۸ شیر

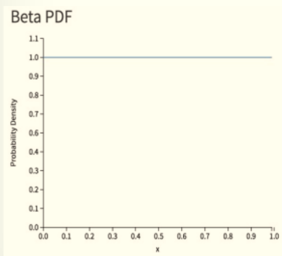
$f_{X\,|(8 \text{ heads},\ 4 \text{ tails})}$



$X \sim \text{Beta}(9{,}5)$

هیچ پرتاب

$f_{X\,|(0 \text{ heads},\ 0 \text{ tails})}$



$X \sim \text{Beta}(1{,}1)$

توزیع مزدوج (Conjugate Distribution)

- دیدیم که توزیع یکنواخت $U(0{,}1)$ در واقع یک توزیع بتا است:

$$U(0{,}1) \sim \text{Beta}(1{,}1)$$

- بنابراین توزیع پیشین X (پیش از $n+m$ پرتاب سکه) توزیع بتا است: $\text{Beta}(1{,}1)$
- توزیع پسین X (پس از $n+m$ پرتاب سکه) نیز توزیع بتا است: $\text{Beta}(n+1, m+1)$
- در این حالت می‌گوییم بتا یک توزیع مزدوج برای توزیع بتا است.
  - شکل پارامتری توزیع پیشین و پسین یکسان است.
  - توزیع بتا همچنین توزیع مزدوج دو توزیع برنولی و دوجمله‌ای است.
  - در عمل توزیع مزدوج بتا به معنی به روز رسانی ساده توزیع احتمال است:
    - کافی است تعداد شیرها و خطاها را به پارامترهای توزیع بتا اضافه کنیم.

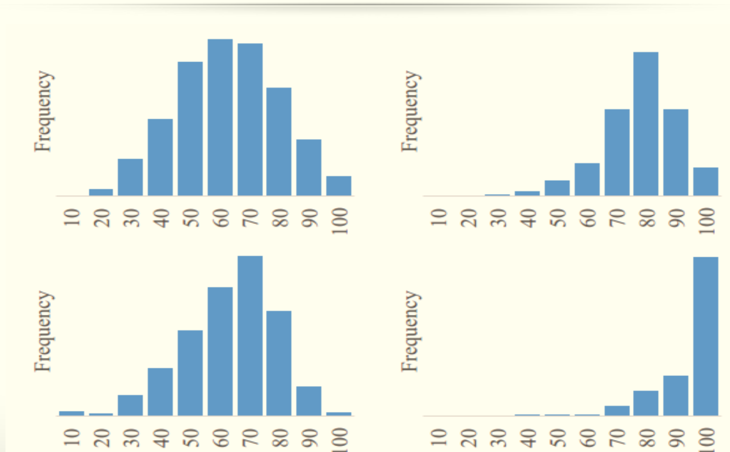
مراحل تخمین بیزی

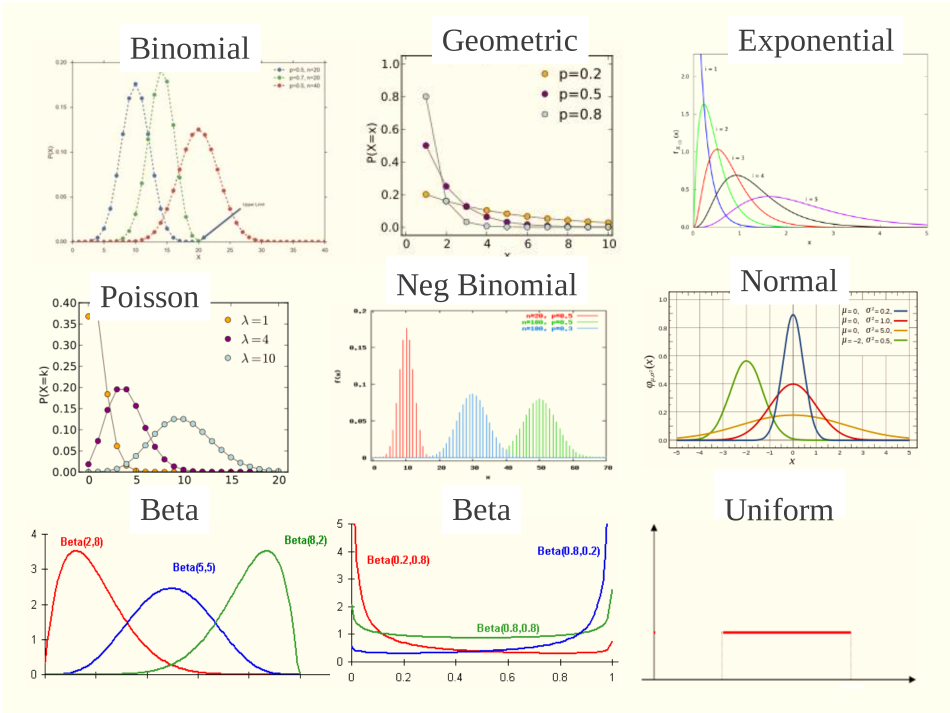
- ابتدا توزیع $Beta(a,b)$ را به عنوان توزیع احتمال پیشین در نظر می‌گیریم.
- سپس $n+m$ بار سکه را پرتاب می‌کنیم، که n تعداد دفعاتی است که شیر آمده است.
- توزیع احتمال پسین X به سادگی و به شکل زیر به روز می‌شود:

$$X|N=n \ \sim \ Beta(a+n, b+m)$$

- اگر هیچ باور و تعصب پیشینی نسبت به احتمال شیر آمدن سکه نداشته باشیم، غالبا توزیع پیشین را $Beta(1,1)$ یا همان توزیع یکنواخت می‌گیریم.

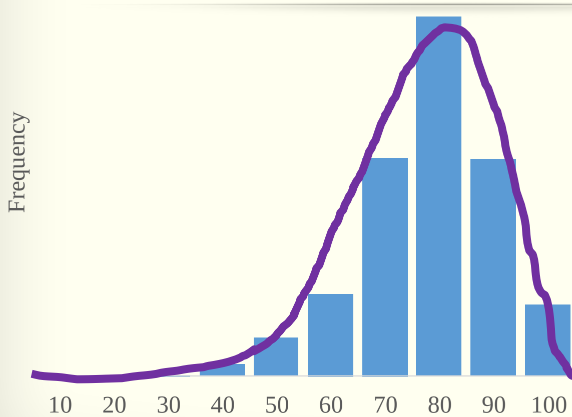
توزیع نمرات تمرین‌ها





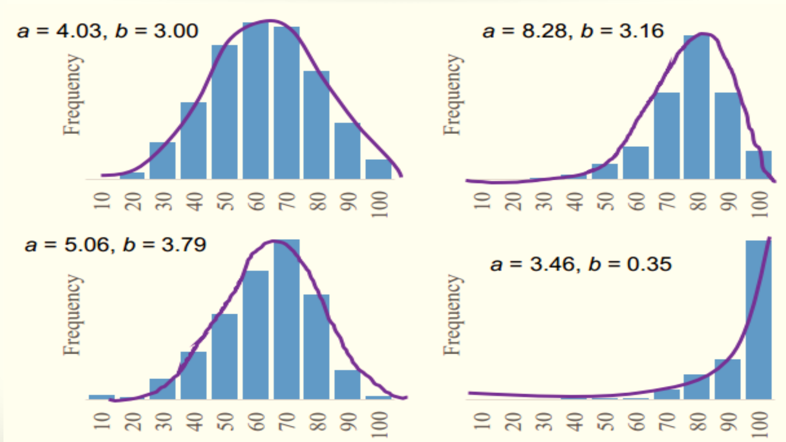
مدل‌سازی برای توزیع نمرات

- توجه کنید که توزیع نمرات تحت هر شرایطی محدود به بازه [0,100] است.
- توزیع بتا بر خلاف توزیع نرمال محدود است و در نتیجه برای مدلسازی نمرات مناسبتر است.

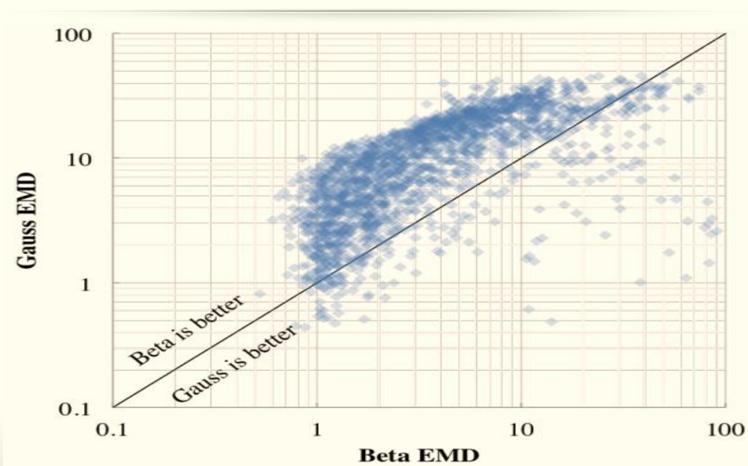


$X \sim \text{Beta}\,(a = 8.28, b = 3.16)$

مدل‌سازی برای توزیع نمرات



مقایسه با توزیع گاوسی



مثالی از فرایند بروزرسانی بیزی