به نام خدا

متغیرهای تصادفی

- گسسته ← تابع جرم احتمال $P_X(x)=P_r\{X=x\}$; Pmf
- پیوسته ← تابع توزیع احتمال PDF
- ترکیبی

$$F_X(x)=P_r\{X\leq x\}$$

معرفی برخی خصوصیات تابع PDF

1) $\forall x \quad 0\leq F_X(x)\leq 1$

2,3) $F_X(-\infty)=0 \quad , \quad F_X(\infty)=1$

4) $F_X(x)$ غیرنزولی

5) $P_r\{X>x\}=1-F_X(x)$

6) $P_r\{x_1 < X \le x_2\} = F_X(x_2) - F_X(x_1)$

به عبارت دیگر با کمک $F_X(x)$ می‌توان احتمال پیش‌آمدهای به فرم $\{x_1 < X \le x_2\}$ را به دست آورد

$\{x_1 < X \le x_2\} = A$

$\{X \in A\}$

$$\{X \le x_2\} = \underbrace{\{X \le x_1\}} \cup \underbrace{\{x_1 < X \le x_2\}}$$

دو پیشامد جدا از هم

$$\Longrightarrow \underbrace{P_r\{X \le x_2\}}_{F_x(x_2)} = \underbrace{P_r\{X \le x_1\}}_{F_x(x_1)} + P_r\{x_1 < X \le x_2\}$$

$$\Longrightarrow P_r\{x_1 < X \le x_2\} = F_x(x_2) - F_x(x_1)$$

7) $$P_r\{X = x_0\} = F_X(x_0) - F_X(x_0^-)$$

$$= \begin{cases} 0 & \text{اگر } F_X(x) \text{ در نقطه } x = x_0 \text{ پیوسته باشد} \\ P_0 & \text{اگر } F_X(x) \text{ دارای پرشی به میزان } P_0 \text{ در نقطه } x = x_0 \text{ باشد} \end{cases}$$

که کران

$$F_X(x_0^-) = \lim_{\substack{\epsilon \to 0 \\ \epsilon > 0}} F_X(x_0 - \epsilon)$$

برای نشان دادن این خصوصیت از خصوصیت قبلی و تعریف حد کمک می‌گیریم.

می‌دانیم که
$$P_r\{x_1 < X \leq x_2\} = F_X(x_2) - F_X(x_1) \qquad (1)$$

اگر قرار دهیم $x_2 = x_0$ و $x_1 \to x_2$ خواهیم داشت

$$\xRightarrow[\substack{x_2 = x_0 \\ x_1 \to x_2}]{(1)} \underbrace{\lim_{\substack{x_1 \to x_2 \\ x_2 = x_0}} P_r\{x_1 < X \leq x_2\}} \quad P_r\{X = x_0\} = \lim_{\substack{x_1 \to x_2 \\ x_2 = x_0}} F_X(x_2) - \lim_{\substack{x_1 \to x_2 \\ x_2 = x_0}} F_X(x_1)$$

$$\Rightarrow P_r\{X = x_0\} = F_X(x_0) - F_X(x_0^-)$$

8) $F_X(x_0^+) = F_X(x_0)$

تابع توزیع احتمال همواره از سمت راست پیوسته است.

با توجه به تعریف تابع توزیع احتمال، مفهوم حد، واضح است

$$F_X(x_0^+) = \lim_{\substack{\epsilon \to 0 \\ \epsilon > 0}} F_X(x_0 + \epsilon)$$

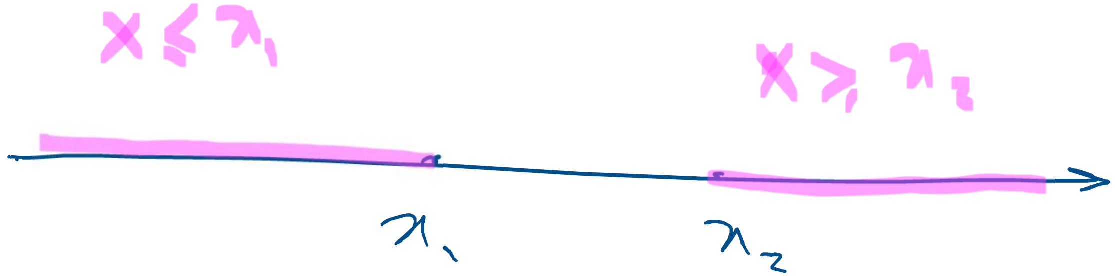
$$F_X(x_0) = P_r\{X \leq x_0\}$$

مثلا با داشتن $F_X(x)$ می‌توانیم احتمال هر پیش‌آمدی را در ارتباط با X حساب کنیم

$$P_r\{x_1 \le X \le x_2\} = \underline{P_r\{X = x_1\}} + \underline{P_r\{x_1 < X \le x_2\}}$$

$$P_r\{X \le x_i\} = F_X(x_i)$$

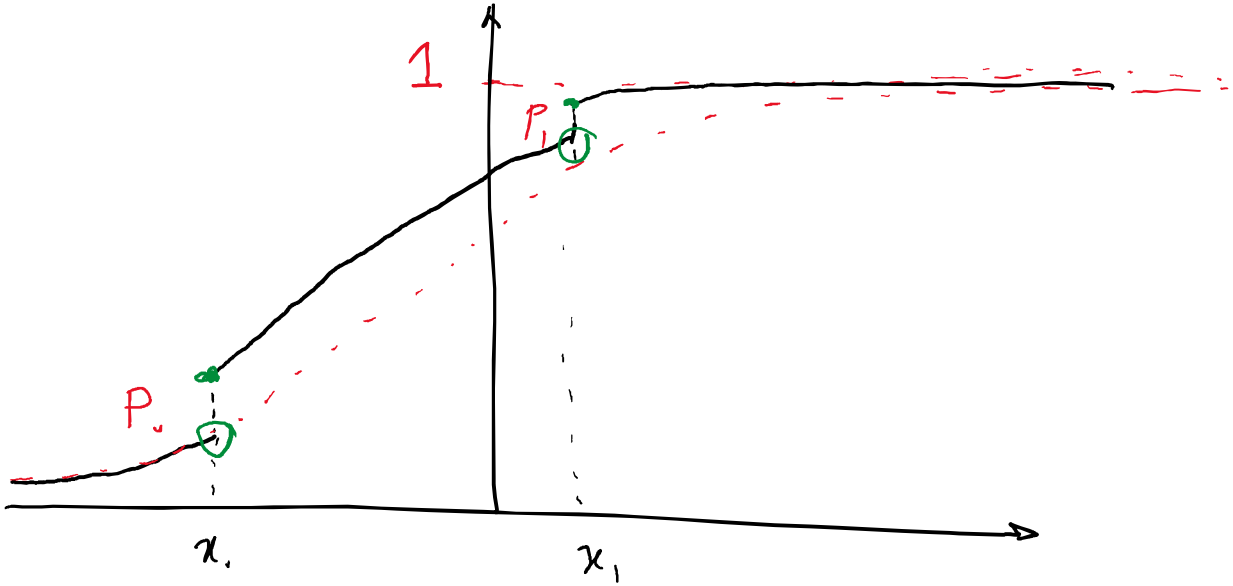
$$P_r\{X > x_i\} = 1 - F_X(x_i)$$



یادآوری: تابع $g(x)$ در نقطه $x = x_0$ پیوسته می‌گوییم اگر تابع $g(x)$ در این نقطه هم از سمت راست و هم از سمت چپ پیوسته باشد یعنی

$$g(x_0^-) = g(x_0^+) = g(x_0)$$

$\Leftarrow$ تابع توزیع احتمال همواره از سمت راست پیوسته است و اگر ناپیوستگی در این تابع وجود داشته باشد، ناپیوستگی از سمت چپ است.



$$P_r\{X=x\} = 0$$

$$x \neq x_0, x_1$$

$$P_r\{X = x_0\} = F_X(x_0) - F_X(x_0^-) = P_0$$

$$P_r\{X = x_1\} = F_X(x_1) - F_X(x_1^-) = P_1$$

* در متغیرهای تصادفی پیوسته ($F_X(x)$ پیوسته است) احتمال نقاط برابر صفر است.

$$F_X(x) = P_r\{X \leq x\}$$

* در متغیرهای تصادفی گسسته، $F_X(x)$ فرم پلکانی دارد (CDF)

$$F_X(x) = \sum_i P_X(x_i)\, u(x - x_i)$$

* مفهوم صدک ها Percentiles

در توابع توزیع احتمال ممکن است ، مقادیری که به ازای آنها احتمال $\{X \leq x\}$ یک مقدار خاص مورد نظر است ، اهمیت داشته باشد یعنی اگر داشته باشیم ،

$$F_X(x) = P_r\{X \leq x\} = u \,, \quad (0 \leq u \leq 1)$$

آنگاه $x = F_X^{-1}(u)$ (نقطه‌ای که $P_r\{X \leq x\} = u$ شده است) برای

ما اهمیت دارد و به آن صدک $100u$ می‌گوییم.

به عنوان مثال:

$$F_X(x) = P_r\{X \leq x\} = 0.1$$

آنگاه $x = F_X^{-1}(0.1)$ را صدک دهم می‌گوییم.

صدک 25 ام را چارک اول نیز می‌گویند. (Quartile)

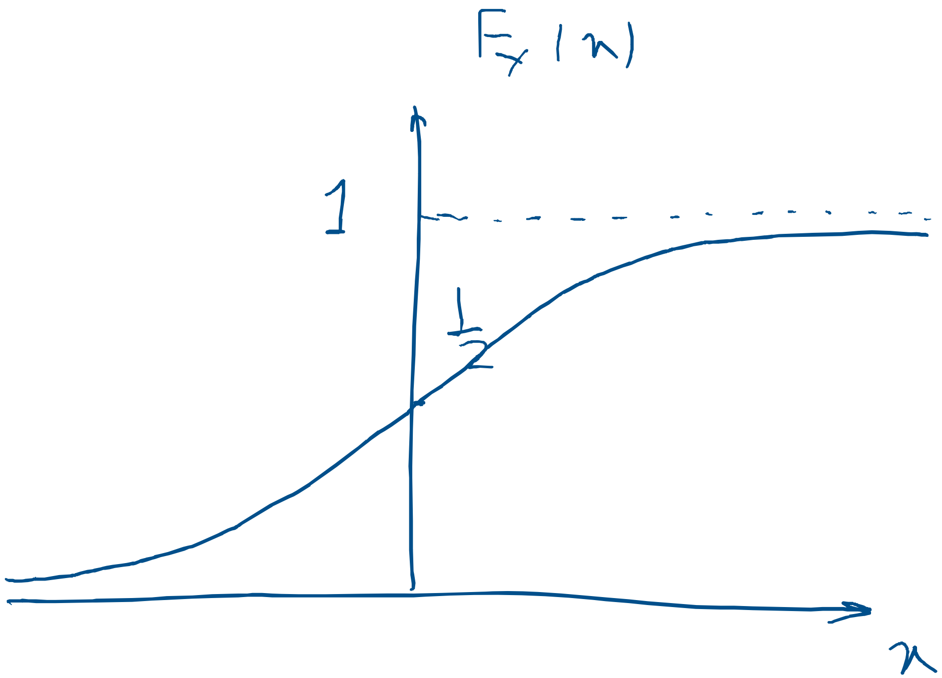
صدک دهم را دهک اول نیز می‌گویند. (Dectile)

صدک پنجاهم (یعنی نقطه ای که در آن $F_X(x) = 0.5$) یک نقطه پراهمیت در تابع توزیع

احتمال است که به آن میانه (median) گفته می‌شود.

به طور مثال:

برای تابع توزیع احتمال شکل مقابل



میانه برابر

$$\text{median} = 0$$

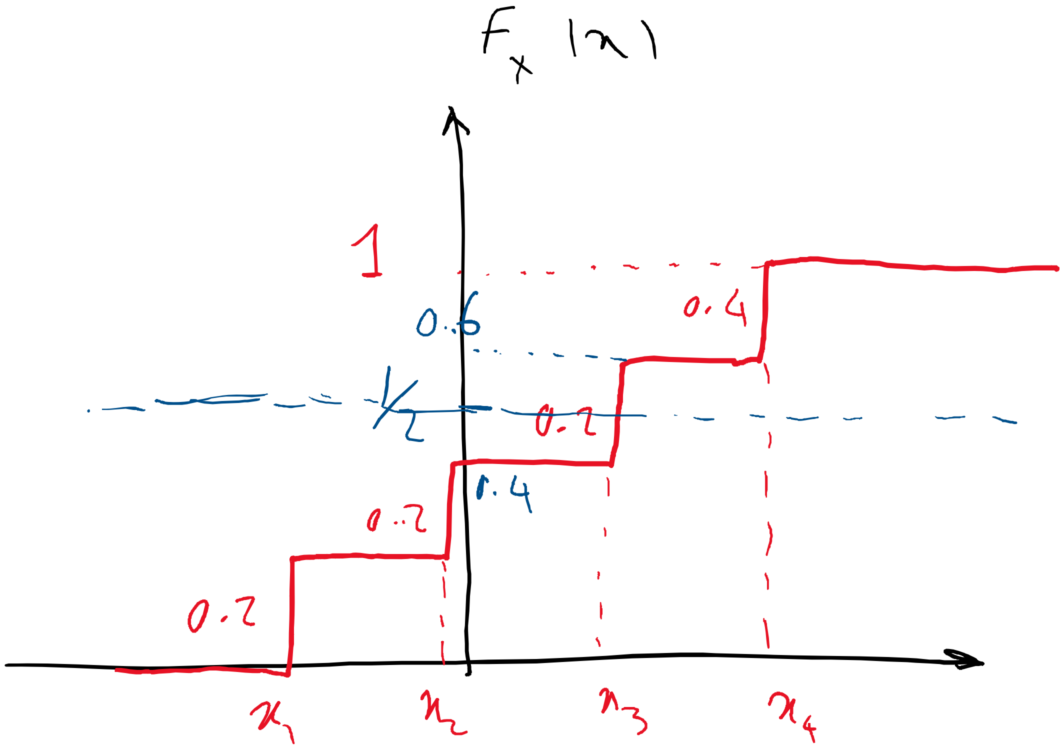
$$F_X(x)\Big|_{x=0} = \frac{1}{2}$$

پیدا کردن میانه برای متغیرهای تصادفی پیوسته که ساده است داریم

$$F_X(x) = \frac{1}{2} \implies \text{median} = F_X^{-1}\left(\frac{1}{2}\right)$$

اما در مورد متغیرهای تصادفی گسسته که تابع توزیع احتمال آنها فرم پلکانی دارد (تابع $F_X(x)$ پیوسته نیست) برای پیدا کردن میانه لازم است با دقت بیشتری به مسأله نگاه کنیم و تعریف دقیق‌تری از میانه ارائه کنیم که هم در حالت پیوسته و هم در حالت گسسته قابل استفاده باشد.

به طور مثال :



$Median = x_3$

در این تابع توزیع احتمال در هیچ نقطه‌ای $F_X(x)$ برابر $\frac{1}{2}$ برقرار نیست.

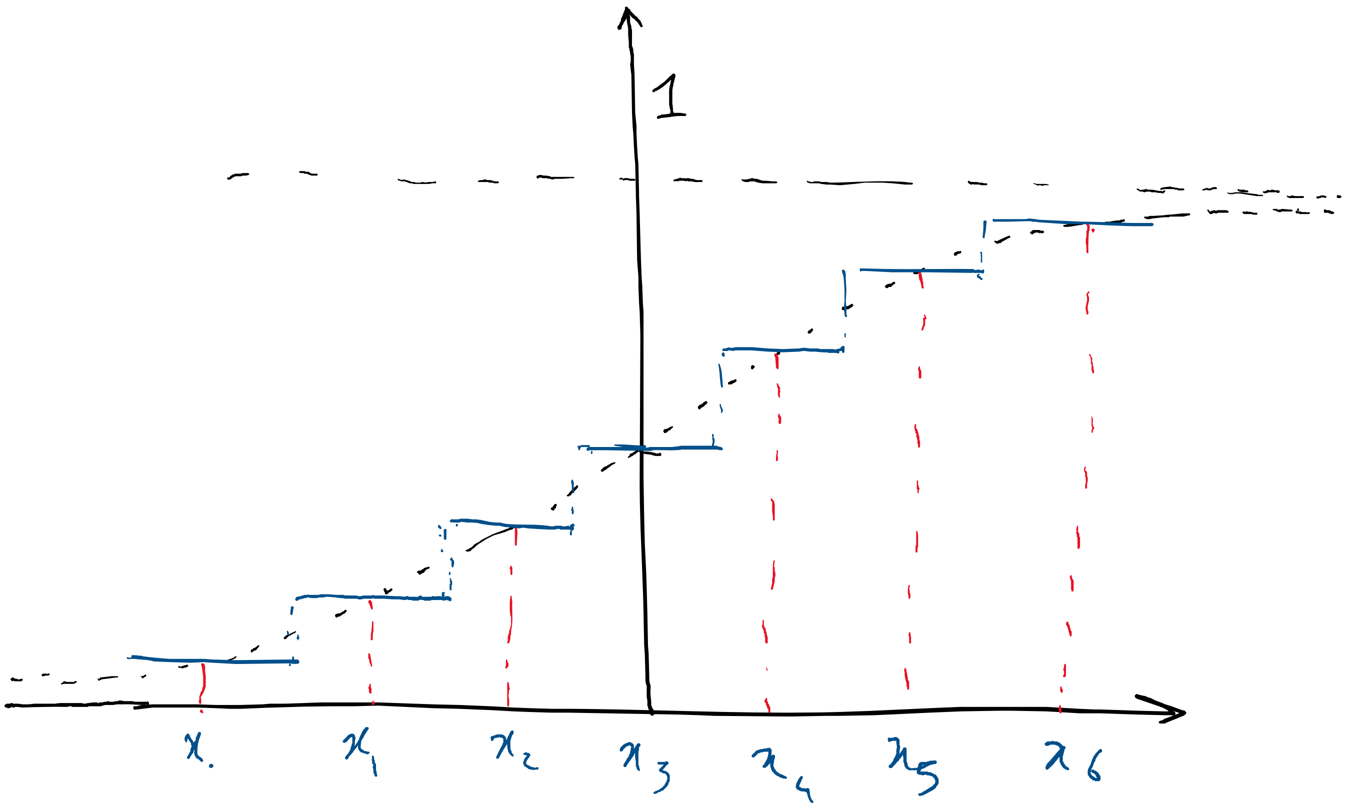
بنا بر این تعریف کلی برای میانه در یک تابع توزیع احتمال $F_X(x)$ به صورت زیر است

$$\text{Median of } F_X(x) = m \iff \begin{cases} P_r\{X \leqslant x\} \leqslant \frac{1}{2} \; ; x \leqslant m \\ P_r\{X \leqslant x\} \geqslant \frac{1}{2} \; ; x \geqslant m \end{cases}$$

* به دست آوردن تابع توزیع احتمال به صورت تجربی (عددی)

در کتاب‌های عملی، ممکن است که در مورد یک متغیر تصادفی X، اندازه‌گیری‌ها یا مشاهداتی را در اختیار داشته باشیم. در این صورت با داشتن این اطلاعات، می‌توانیم

تابع توزیع احتمال متغیر تصادفی را به صورت تجربی (یا تقریبی یا عددی) به دست بیاوریم.



$$F_X(x) = P_r\{X \le x\}$$

$$F_X(x) = P_r\{X \le x\} \cong \frac{n_x}{n} \qquad (1)$$

کل تعداد مشاهدات یا اندازه‌گیری‌ها

که در آن n_x تعداد حالت‌هایی است که برای آنها شرط $\{X \le x\}$ برقرار است.

به عبارت دیگر برای به دست آوردن $F_X(x)$ به صورت تجربی، ابتدا باید اندازه‌گیری‌ها یا مشاهدات را مرتب کنیم و سپس برای هر نقطه x، از عبارت (1)، مقدار تقریبی (تجربی) $F_X(x)$ را به دست بیاوریم و نمودار پله‌ای آن را رسم کنیم.

هر چه تعداد مشاهدات یا اندازه‌گیری‌ها را بیشتر کنیم، به تعریف نزدیک‌تری از تابع توزیع

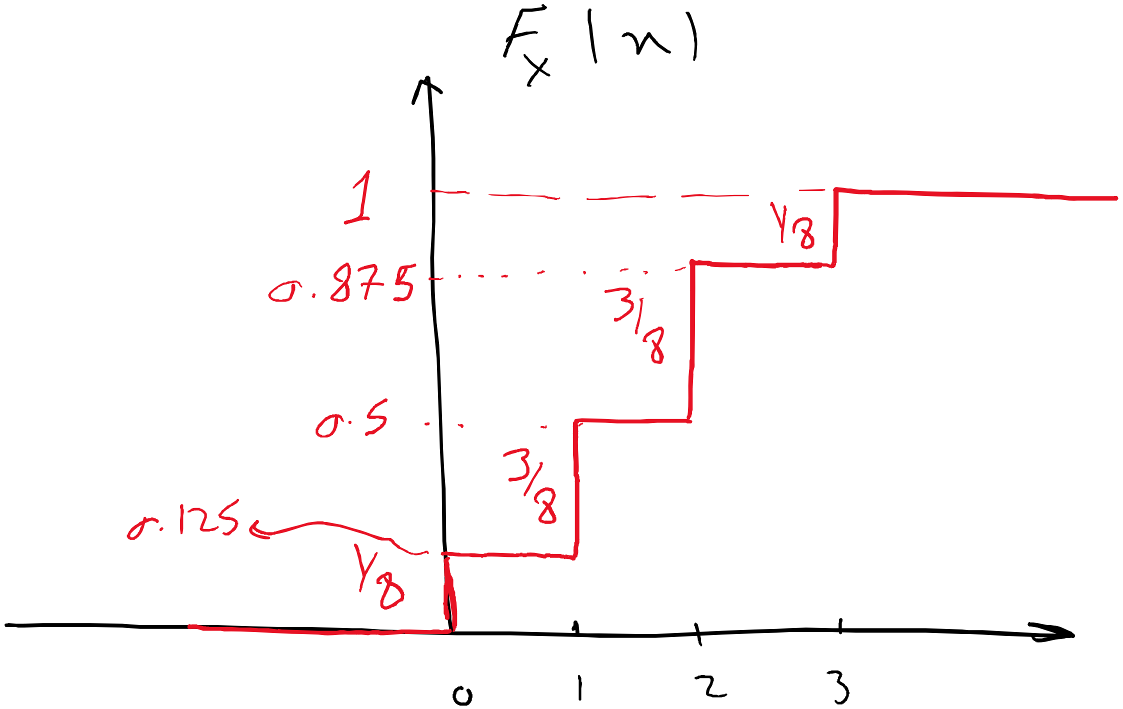
احتمال متغیر تصادفی مورد نظر خواهیم رسید.

مثال - یک سکه سالم را سه بار پرتاب کنیم، فرض می کنیم متغیر تصادفی X نشان دهنده تعداد شیرها در سه بار پرتاب سکه باشد. می دانیم که تابع توزیع احتمال X ، یک تابع توزیع احتمال درجه ای به فرم زیراست

$$F_X(x) = \sum_{i=0}^{3} P_X(i)\, u(x-i)$$

که در آن

$$P_X(i) = P_r\{X=i\} = \binom{3}{i}\left(\frac{1}{2}\right)^{i}\left(\frac{1}{2}\right)^{3-i} = \binom{3}{i}\left(\frac{1}{2}\right)^{3}$$



این آزمایش را بصورت تجربی، ۱۰ بار تکرار کرده ایم و به نتایج زیر رسیده ایم. تابع توزیع احتمال تقریبی متغیر تصادفی را بدست بیاوریم.

x	0	1	2	3
ℓ	2	4	3	1

$$F_X(x) = P_r\{X \leq x\} \cong \frac{n_x}{n}$$

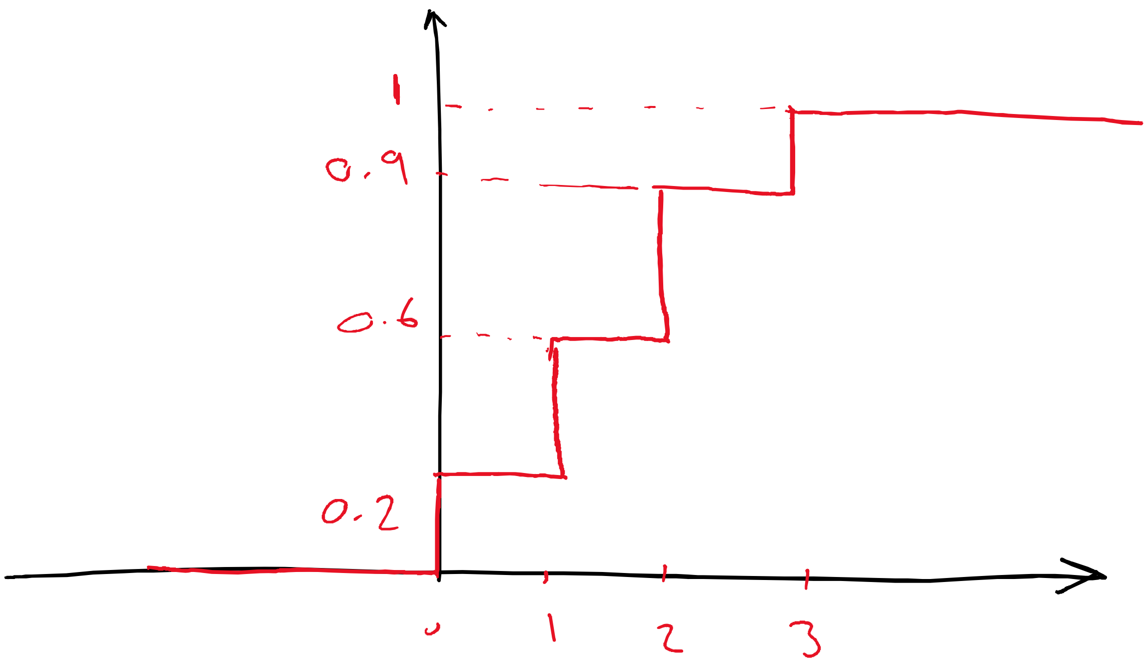
در این آزمایش $n=10$

$$n_0 = 2 \implies F_X(x) = P_r\{X \leq x\}\Big|_{x=0} \cong \frac{n_0}{n} = 0.2$$

$$n_1 = 2+4 \implies F_X(x) = P_r\{X \leq x\}\Big|_{x=1} \cong \frac{n_1}{n} = 0.6$$

$$n_2 = 2+4+3 \implies F_X(x) = P_r\{X \leq x\}\Big|_{x=2} \simeq \frac{n_2}{n} = 0.9$$

$$n_3 = 2+4+3+1 \implies F_X(x) = P_r\{X \leq x\}\Big|_{x=3} \simeq \frac{n_3}{n} = 1$$



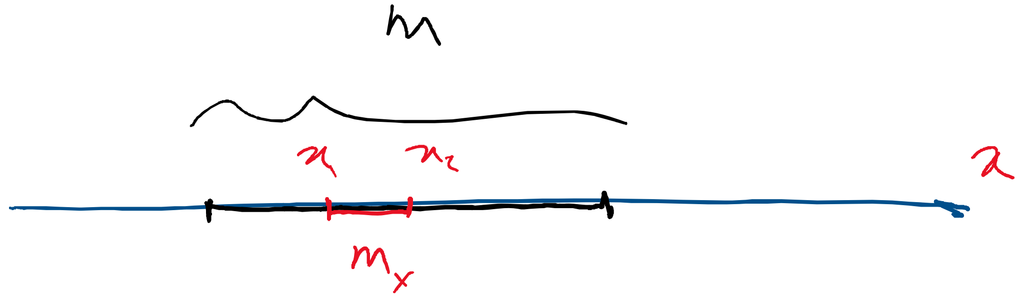
در ادامه می‌خواهیم توابع دیگری برای تحلیل متغیرهای تصادفی معرفی کنیم.

- تابع چگالی احتمال Probability density function (pdf)

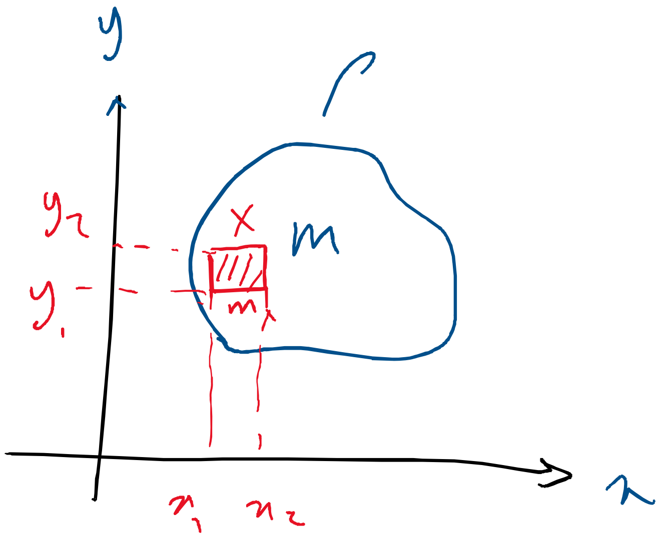
در کتاب متغیرهای تصادفی به دنبال تابعی هستیم که با استفاده از آن و با انتگرال‌گیری از آن بر روی ناحیه‌های مرتبط با پیشامدهای متغیر تصادفی، بتوانیم به احتمال آن پیشامد برسیم. به همین دلیل این تابع شباهت زیادی با توابع چگالی جرمی دارد و از این جهت به آن تابع چگالی احتمال می‌گوییم.

تابع چگالی جرمی

ρ



$$\int_{x_1}^{x_2} \rho \, dx$$



$$m_x = \int_{y_1}^{y_2} \int_{x_1}^{x_2} \rho \, dx \, dy$$

با توجه به اینکه می‌دانیم با داشتن $F_X(x)$ می‌توانیم احتمال هر پیشامدی در راستای X (فضای یک بعدی) را به دست بیاوریم، می‌توانیم تابع چگالی احتمال متغیر تصادفی X را به صورت زیر تعریف کنیم،

$$\underbrace{f_X(x)}_{\text{تابع چگالی احتمال متغیر تصادفی } X} \triangleq \frac{d}{dx} F_X(x)$$

$$\Longrightarrow \quad F_X(x) = \int_{-\infty}^{x} f_X(\alpha)\, d\alpha + C$$

$$F_X(-\infty) = 0$$ از طرف دیگر می‌دانیم که

$$\Longrightarrow \quad \underbrace{F_X(x)\Big|_{x=-\infty}}_{0} = \underbrace{\int_{-\infty}^{-\infty} f_X(\alpha)\, d\alpha}_{0} + C$$

$\Rightarrow \quad C = 0$

$$\Rightarrow \begin{cases} f_X(x) = \dfrac{d}{dx} F_X(x) \qquad (1 \\ \text{و} \\ F_X(x) = \displaystyle\int_{-\infty}^{x} f_X(\alpha)\, d\alpha \end{cases}$$

یعنی یک رابطه یک به یک بین $F_X(x)$ و $f_X(x)$ وجود دارد.

به دلیل اینکه رابطه یک به یک بین $F_X(x)$ و $f_X(x)$ وجود دارد، هر خصوصیتی که برای

تابع توزیع $F_X(x)$ بیان کردیم به طور مشابه برای تابع چگالی $f_X(x)$ نیز

برقرار است.

2) $\int_{-\infty}^{\infty} f_X(\alpha)\, d\alpha = 1$

تابع چگالی احتمال $f_X(x)$ یک تابع نرمالیزه است.

$$F_X(x) = \int_{-\infty}^{x} f_X(\alpha)\, d\alpha \xrightarrow{x \to \infty} \underbrace{F_X(\infty)}_{1} = \int_{-\infty}^{\infty} f_X(\alpha)\, d\alpha$$

$$\Rightarrow \int_{-\infty}^{\infty} f_X(\alpha)\, d\alpha = 1$$

3) $\forall x;\ f_X(x) \geq 0$

زیرا $f_X(x)$ مشتق تابع $F_X(x)$ است که می‌دانیم $F_X(x)$ غیرنزولی است.

بنابراین $f_X(x)$ همواره غیرمنفی است.

4) $P_r\{x_1 < X \leq x_2\} = F_X(x_2) - F_X(x_1) = \int_{x_1}^{x_2} f_X(\alpha)\, d\alpha$

$$F_X(x_2) = \int_{-\infty}^{x_2} f_X(\alpha)\, d\alpha$$

$$F_X(x_1) = \int_{-\infty}^{x_1} f_X(\alpha)\, d\alpha$$

$\Leftarrow$ احتمال هر پیش آمدی به فرم $\{x_1 < X \leq x_2\}$ را می‌توان با کمک $f_X(x)$ به دست آورد.

5) $$P_r\{X=x_0\} = F_X(x_0) - F_X(x_0^-) = \int_{x_0^-}^{x_0} f_X(\alpha)\,d\alpha$$

$$= \begin{cases} 0 & \text{اگر } f_X(x) \text{ در نقطه } x=x_0 \text{ پیوسته باشد} \\ P_0 & \text{اگر } f_X(x) \text{ در نقطه } x=x_0 \text{ ضربه‌ای به وزن } P_0 \text{ داشته باشد} \end{cases}$$

$\Leftarrow$ با استفاده از $f_X(x)$ می‌توان احتمال هر پیشامد در ارتباط با X را محاسبه کرد.

6) $\Pr\{X \in A\} = \int_A f_X(\alpha)\, d\alpha$

$A \subseteq \mathbb{R}$