بسم الله الرحمن الرحیم

جلسه سوم ۱۴۰۰/۷/۲۰

Probability → Computational approach

احتمال ← تابع چگالی احتمال

احتمال توزیع تابع

Probability Distribution function ← فراوانی

Correlation ← همبستگی‌ها را بررسی کند

نگاه کردن جدا جدا تک تک سلول که مستقل از آرایش سایر سلول‌ها

one-Point PDf

1	1	2	3	2	4	5	2

$n(2) = 3$
$n(1) = 2$
$n(3) = 1$
$n(4) = 1$
$n(5) = 1$

فراوانی

1	2	1	2	3	5	2	4

$n(1) = 2$
$n(2) = 3$
$n(3) = 1$
$n(4) = 1$
$n(5) = 1$

فراوانی هم برابر هستند

فراوانی → Probability

تک نقطه one-Point Probability

از حیث ترتیب قرار گرفتن اعداد در کنار هم متفاوت هستند

Correlation → متفاوت هستند

نظر به آرایش و چگونگی قرار گرفتن اعضا است

Relative Frequency

$$RF = \frac{n}{N}$$

n ← تعداد حالت‌های مطلوب

N ← تعداد کل رخدادها

$$\{\xi\} : \{\xi_1, \xi_2, \dots \quad \xi_N\}$$

تعداد کل تکرار یک آزمون / آزمایش یا اندازه‌گیری / شبیه‌سازی

$$RF(\xi = x) = \frac{n(\xi = x)}{N}$$

$$\text{Probability} = \lim_{N\to\infty} RF \qquad \boxed{\text{long-Run}}$$

$$P_\xi(x) = \lim_{N\to\infty} RF_\xi(x) = \frac{\text{Abundance}(x)}{\#\text{ of all events}}$$

ترتیب و آرایش قرار گرفتن اعداد در شکل $P_\xi(x)$ نقشی ندارد } One-Point Statistics

آماره تک نقطه‌ای است

Probability Density function (PDF)

$$P_\xi(x) = \lim_{N\to\infty} RF_\xi(x) \equiv \underbrace{p_\xi(x)}_{\text{PDF}} \, \Delta x$$

$$x \leq \xi < x + \Delta x$$

احتمال بین مقدار

$$\xi \in [x, x+\Delta x) \;, \quad \boxed{x - \frac{\Delta x}{2} < \xi < x + \frac{\Delta x}{2}}$$

$$p_\xi(x) = \frac{P_\xi(x)}{\boxed{\Delta x}} \simeq \frac{n(x)}{N\boxed{\Delta x}}$$

اصل ← بهنجارش

$$\sum_{i=1}^{\boxed{M}} \underbrace{p_\xi(x_i)} = \sum_{i=1}^{\boxed{M}} \frac{n_\xi(x_i)}{N} = \frac{1}{N} \underbrace{\sum_{i=1}^{M} n_\xi(x_i)} = 1$$

تعداد کل رخدادها = N

$$\int \underbrace{p_\xi(x)}_{\text{PDF}} dx = 1$$

N تعداد اعضای مجموعه (رخداد) ← دانشجویان

M تعداد کلاس‌ها ←

$N \geq M$

- سال اول
- سال دوم
- سال سوم

$$\int p_\xi(x)\, dx = A \longrightarrow p_\xi(x) \longrightarrow \bar{p}_\xi(x) = \frac{p_\xi(x)}{A}$$

روش رسم بهنجار شکل

* Mathematical Description of "PDF"

$$p_\xi(x) = \left\langle \delta_D(x-\xi) \right\rangle_\xi \longleftarrow$$

↓

Dirac Delta function

Recall $\underbrace{\langle F \rangle} = \int d\Gamma\, F\, \underbrace{\rho(\Gamma)}$

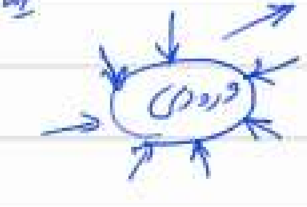
$$\hookrightarrow p_\xi(x) = \left\langle \delta_D(x-\xi) \right\rangle = \int d\xi\, \delta_D(x-\xi)\, p_\xi(\xi)$$

$$= p_\xi(x)$$

* Computational approach

$\{\xi\}:\{\xi_1,\xi_2,\dots,\xi_N\}$ $N\to\infty$

$M \equiv \#$ of classes $\underbrace{\Delta x}_{\text{بازه}} \equiv \dfrac{\overbrace{\max(\xi)}^{\text{بیشینه}}-\overbrace{\min(\xi)}^{\text{کمینه}}}{M}$

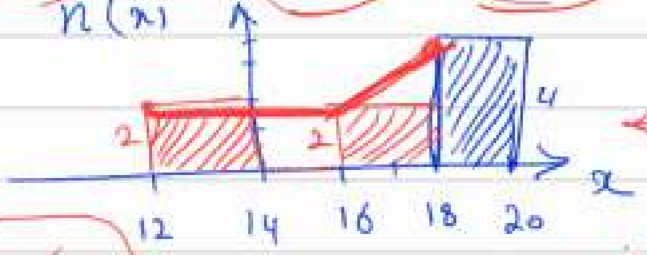


$\underline{x},\ \underline{x+\Delta x}$

هر چقدر که M بزرگتر شود شکل تابع توزیع دارای افت و خیز بیشتری خواهد بود و به طور دقیق مشخص نمی‌کند که دقیقا چگونه است؟

$\xi:\{12, 16, 17, 18.5, 19.5, 18.25, 18.75, 12.8\}$

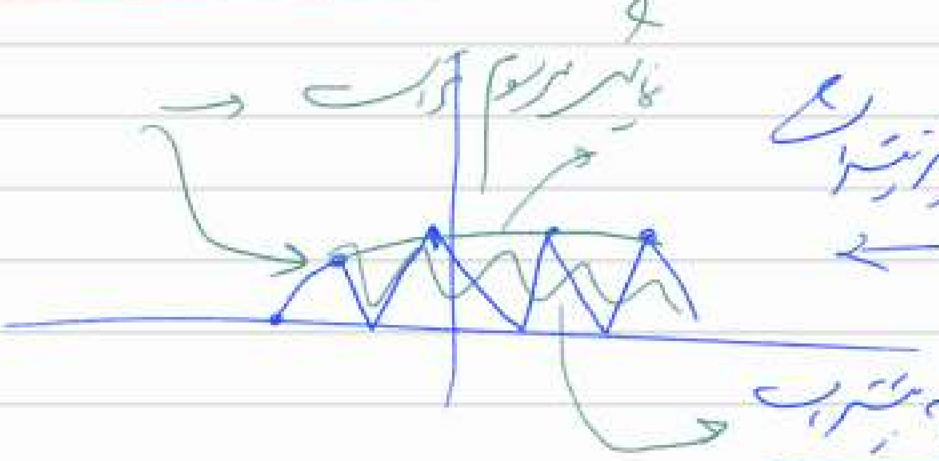
$\Delta x = 2$



$\{\xi\}$

$\to\{\underbrace{12,12}_{2},\underbrace{16,16}_{2},\underbrace{18,18,18,18}_{4}\}$

در عوض دید دادیم



$\boxed{\Delta x}$

افت و خیز بیشتر

دقت بیش از

حد مجموعه اولیه بیشتر بشه

Weight function تابع وزن

PDF
- Ⓐ Weight function
- Ⓑ : Kernel Estimation

روی کدام محاسبات این تقسیم بندی را انجام میدهیم

Ⓐ Weight function

① Simple estimation

Weight funct

```
Import Data {ξ}.{ξ_1 --- ξ_N}
$\Delta x = \frac{Max(\xi) - Min(\xi)}{M}$   M ← input variable
loop i=1, N
    Index → $k = \frac{\xi(i)}{\Delta x}$
    $p(k) = p(k) + 1$
End loop
$p = \frac{p}{N}$
Write $k \times \Delta x$ , $p(k)$
```

شروع میکنیم از داده ها اگر میرسیم (k) متعلق به کدام کلاس است بعد آن کلاس را پیدا کرده و تعداد نقاط کلاس را افزایش میدهیم

عددی نشان دهنده کلاس

$k = 16$

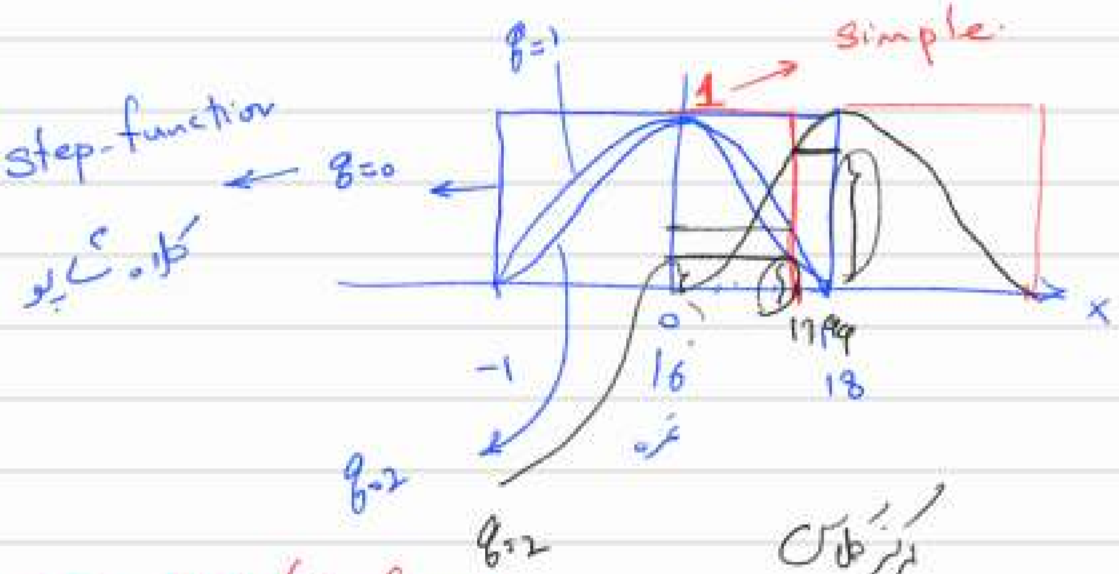
$\boxed{\Delta x} = 2$

$16, \boxed{17.99}$ → 18

[illegible]

یکی از انواع معروف تابع وزن است

$$W(x, q) = \frac{(1-x^2)^q}{2^{2q+1}\,\beta(q+1, q+1)}\ \mathbb{1}\{|x| \leq 1\}$$



q=1
simple
1
step-function
q=0
کلاه شاپو
-1
0
16
17
18
x
q=2
q=2
مرکز کلاس

② Weight function

$$\rightarrow \boxed{p(k) = p(k) + \underbrace{W(k\Delta x, \xi_i)}_{1}} \leftarrow \boxed{p(k) = p(k)+1}$$

$$\boxed{W(k\Delta x, \xi_i) = \frac{\left(1-(k\Delta x-\xi_i)^2\right)^q}{2^{2q+1}\,\beta(q+1, q+1)}}$$

در هر لحظه که داریم تابع فراوانی را حساب می‌کنیم وزن دهی هم انجام می‌دهیم

$$\boxed{\text{A typical Weight function}}$$

[illegible]

تابع وزنی استفاده کرده‌ایم دیگه

Ⓑ Kernel Estimation

فقط مختصر شده PDF ش ← وقتی کوانتم به هر دلیلی تابع نمودار اندازه‌گیری

Smooth ← نرم کنیم

Convolution (پیچش)

تابع جدیدی می‌سازم ← $$F(x) = f(x) \otimes K(x)$$

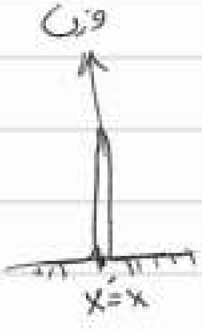
تابع اولیه f با هسته K

$$= \int dx'\, \underbrace{K(x-x')}_{\text{Kernel}}\, f(x')$$

window function

$$F(x) = \sum_{i=1}^{N} K(x-x_i)\, f(x_i)\, \Delta x$$

$$\text{if} \quad K(x-x') = \delta_D(x-x')$$



$$F(x) = \int dx'\, K(x-x')\, f(x')$$

$$F(x) = \int dx'\, \underbrace{\delta_D(x-x')}\, \underbrace{f(x')}$$

$$\underline{F(x) = f(x)}$$

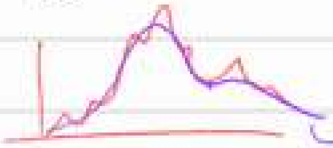
[illegible]

بدون تغییر

با افزایش پهنای Kernel سهم همه بین‌ها در مقدار تابع در نقطه x افزایش می‌یابد.

هموارتر از جعبه PDF

در روش کرنل - ابتدا با یکی از روش‌های ساده p(x) را حساب می‌کنیم. و بعد موضوع نرم‌سازی



Kernel

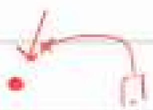
به کار برده می‌شود

اغلب به دلیل سادگی

ایده‌آلی

* مشابه اثر بیم

Beam Effect ← مات شدن تصویر



$$F(x) = \int dx'\, K(x-x')\, f(x')$$

$$F(x) = \int d\lambda\, e^{-i\lambda x}\, \bar{K}(\lambda)\, \bar{f}(\lambda)$$

$$\boxed{\bar{F}(\lambda) = \underset{\uparrow}{\bar{K}(\lambda)}\ \underset{\uparrow}{\bar{f}(\lambda)}}$$

اگر از تبدیل فوریه استفاده کنیم ⟶

$$\begin{tikzcd} K(x) \arrow[r, "\overset{①}{F.T.}"] & \bar{K}(\lambda) \arrow[dr] & \\ f(x) \arrow[r, "\overset{②}{F.T}"] & \bar{f}(\lambda) \arrow[r] & \bar{F}(\lambda) = \bar{K}(\lambda)\bar{f}(\lambda) \arrow[d, "③\ I.F.T"] \\ & & F(x) \end{tikzcd}$$

متناظر هستند

$$\boxed{F(x) = \overset{①}{\int} dx'\, K(x-x')\, f(x')}$$

از مرتبه N^2 > خیلی زودتر

$\{\xi\}.\{\xi_1, \xi_2 \dots \xi_N\}$

N تعداد داده‌ها

M تعداد ستون‌ها

$p_\xi(x) \longrightarrow$ از روش ساده حساب شود

$$\{x_1 \dots x_N\} \; \bigcirc\!\!\!\!\!\!\; p_\xi(x) \longrightarrow \tilde{p}_\xi(x) = \int dx' \, K(x - x') \, p_\xi(x)$$

تعداد ستون‌ها $\to M$

$$\underbrace{p_\xi(x_i)}_{\text{first loop}} = \sum_{j=1}^{M} \Delta x \, \underbrace{K(x_i - x_j)}\,\underbrace{p_\xi(x_j)}$$

first loop → ← Second loop

Algorithm

Import Data $\{\xi\}, \{\xi_1 \dots \xi_N\}$.

$$\Delta x = \frac{\max(\xi) - \min(\xi)}{M \leftarrow \text{input}}$$

```
loop i=1, N
    $R = \frac{\xi_i}{\Delta x}$
    $p(R) = p(R) + 1$   ←   زیاد
End loop
```

$p = \frac{p}{N \Delta x} \rightarrow$ PDF   تابع خطا برای احتمال   الگوریتم ساده

Kernel Estimation

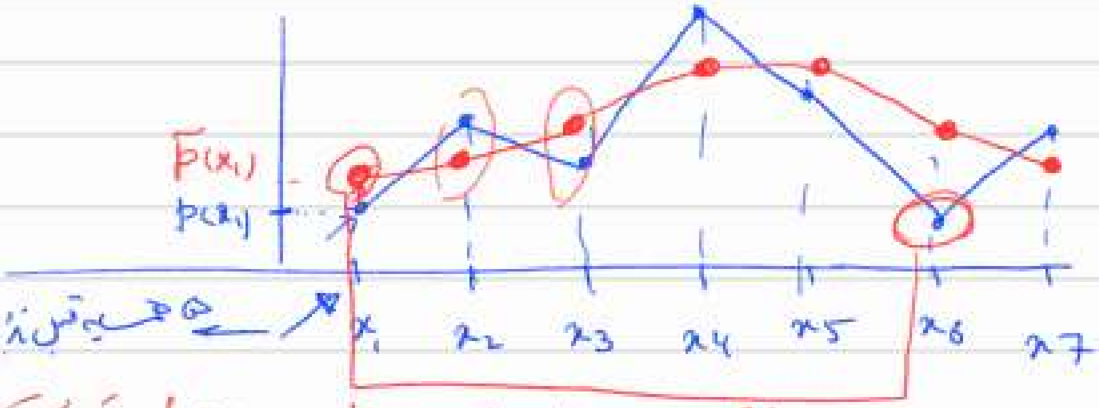
```
loop class $k \leftarrow x$                                  $O(M)$
    loop $k' \leftarrow x' \leftarrow O(M)$
        $\bar{p}(k) = \bar{p}(k) + K(\overset{x}{\widetilde{k\Delta x}} - \overset{x'}{\widetilde{k'\Delta x}})\, p(k)$
    End loop
End loop
```

$O(M^2)$

قابل بهتری

Normalization $\int \bar{p}(x)\, dx = 1$

- $\bar{p}(x)$
- $p(x)$



۵ همسایه نداره

نقش همسایه‌ها عوامل در اصلاح مقدار $p(x_1)$ تغییرات ؟

$K(x - x')$

$n=۵$ تعداد همسایه‌هایی که اثر کنه

n_s

x_{min} x_{max}

سمت مرزی را اصلاح کنم.

```
loop $k = k_{min}, k_{max}$
    loop $k' = k - \min(\overset{5}{n}, k - k_{min})$
             $k + \min(n, k_{max} - k)$
    End loop
```

if $k = k_{min}$ — $k' = k, 5$

$k' = k_{min} - \min(5, k_{min} - k_{min}) = k_{min}$ ✓

$k' = k_{min} + m(5, \cancel{k}_{max} - k_{min}) = k_{min} + 5$

[illegible] صحیح است ✓

if $k = k_{max}$ $\longrightarrow$ $k' = k_{max} - 5, k_{max}$

$k' = k_{max} - \min(5, k_{max} - k_{min}) = k_{max} - 5$

$k' = k_{max} + \min(5, 0) = k_{max}$

تابعی ← p(x) = ✓

رهیافت تحلیلی p(x) به روش ساده ✓

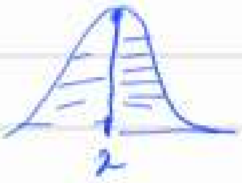
p(x) ← اگر این را هم نشان دادیم.

تاکیدم این بوده که آماره مستقل از نقاطی که می‌بینیم پس مرتبه‌ی آمارات

و ترتیب قرار گرفتن داده‌ها دیگری داده‌ها در این محاسبه نقشی نداشت.

$$\underbrace{P_\xi(x)} = \underbrace{p_\xi(x)\,\Delta x} \longleftarrow \boxed{\lim_{\underline{dx \to 0}} \frac{dP_\xi(x)}{dx} = p_\xi(x)}$$

جا با عوض کرد ←



→ Gaussian

کمی تیزش را ما کم می‌کنیم ← مدل دلخواه.

One-Point Statistics → Joint / Conditional / Correlation

$$\frac{\text{PDF - Transformation}}{\text{MCMC}}$$

↑

از مدل سازی داده ← یکی از موضوعات پایه‌ای شروع می‌کنیم