

شناسایی الگو
فصل هفتم
روش حداکثر انتظار

Expectation Maximization Method

محاسبه احتمال (مثال ۱)

a Maximum likelihood



5 sets, 10 tosses per set

○ احتمال رو آمدن برای دو سکه معلوم نیست. با یک آزمایش قصد داریم آن را تعیین می کنیم.

○ به دلخواه یک سکه را انتخاب نموده (نوع سکه معلوم است) ده بار پرتاب می کنیم. حالت های آمده را یادداشت می کنیم و سپس احتمال هر کدام را می یابیم.

محاسبه احتمال (مثال ۱)

a Maximum likelihood

	H T T T H H T H T H
	H H H H T H H H H H
	H T H H H H H T H H
	H T H T T T H H T T
	T H H H T H H H T H

5 sets, 10 tosses per set

$$\hat{\theta}_A = \frac{24}{24 + 6} = 0.80 \quad \hat{\theta}_B = \frac{9}{9 + 11} = 0.45$$

○ احتمال رو آمدن برای دو سکه معلوم نیست. با یک آزمایش قصد داریم آن را تعیین می کنیم.

○ به دلخواه یک سکه را انتخاب نموده (نوع سکه معلوم است) ده بار پرتاب می کنیم. حالت های آمده را یادداشت می کنیم و سپس احتمال هر کدام را می یابیم.

حداکثر انتظار (مثال ۲)



احتمال رو آمدن برای دو سکه معلوم نیست با یک آزمایش قصد داریم آن را تعیین می کنیم.

به دلخواه یک سکه را انتخاب نموده ده بار پرتاب می کنیم. در اینجا سکه انتخاب شده معلوم نیست اما نتایج ده بار پرتاب آن معلوم است.

حداکثر انتظار (مثال ۲)



هر مجموعه داده داریم:

$$X = (x_1, x_2, x_3, x_4, x_5)$$

$$Z = (z_1, z_2, z_3, z_4, z_5)$$

تعداد رو در مرحله i

نوع سکه انتخاب شده در هر مرحله i

مجهولات:

$$\theta = (\theta_A, \theta_B)$$

احتمال آمدن رو برای سکه A و B

حداکثر انتظار

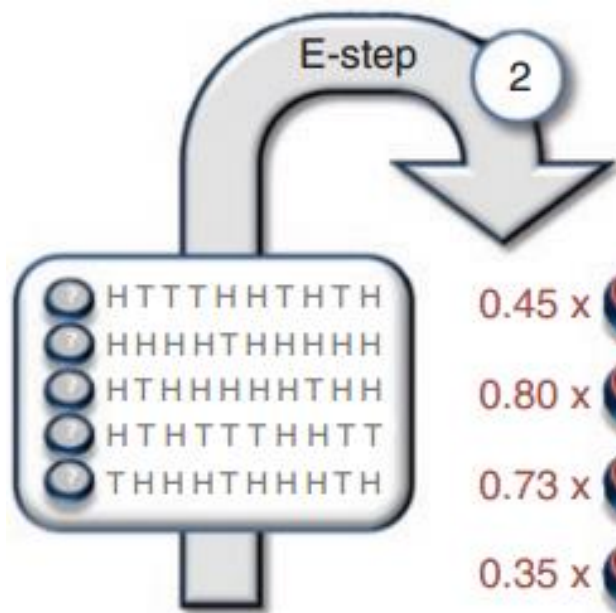


در مثال فوق θ مجهول ماست (پارامترهای نامعلوم).
اما برخی از داده های ورودی (مشاهدات ناقص) Z
موجود نیستند.

EM دو مرحله دارد.

۱- مرحله Expectation

۲- مرحله Maximization

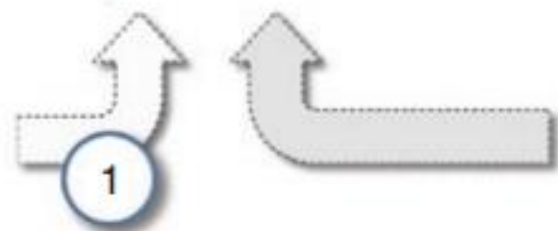


$$\hat{\theta}_A^{(0)} = 0.60$$

$$\hat{\theta}_B^{(0)} = 0.50$$

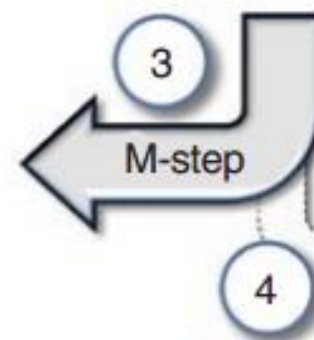
0.45 x		0.55 x	
0.80 x		0.20 x	
0.73 x		0.27 x	
0.35 x		0.65 x	
0.65 x		0.35 x	

Coin A	Coin B
$\approx 2.2 \text{ H}, 2.2 \text{ T}$	$\approx 2.8 \text{ H}, 2.8 \text{ T}$
$\approx 7.2 \text{ H}, 0.8 \text{ T}$	$\approx 1.8 \text{ H}, 0.2 \text{ T}$
$\approx 5.9 \text{ H}, 1.5 \text{ T}$	$\approx 2.1 \text{ H}, 0.5 \text{ T}$
$\approx 1.4 \text{ H}, 2.1 \text{ T}$	$\approx 2.6 \text{ H}, 3.9 \text{ T}$
$\approx 4.5 \text{ H}, 1.9 \text{ T}$	$\approx 2.5 \text{ H}, 1.1 \text{ T}$
$\approx 21.3 \text{ H}, 8.6 \text{ T}$	$\approx 11.7 \text{ H}, 8.4 \text{ T}$



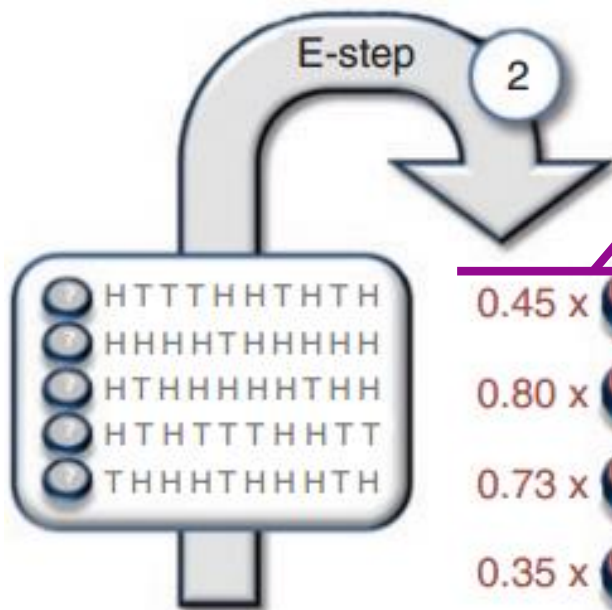
$$\hat{\theta}_A^{(1)} \approx \frac{21.3}{21.3 + 8.6} \approx 0.71$$

$$\hat{\theta}_B^{(1)} \approx \frac{11.7}{11.7 + 8.4} \approx 0.58$$



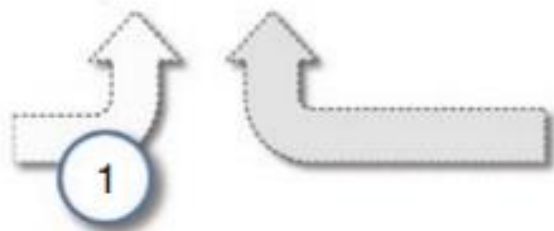
$$\hat{\theta}_A^{(10)} \approx 0.80$$

$$\hat{\theta}_B^{(10)} \approx 0.52$$



$$\hat{\theta}_A^{(0)} = 0.60$$

$$\hat{\theta}_B^{(0)} = 0.50$$



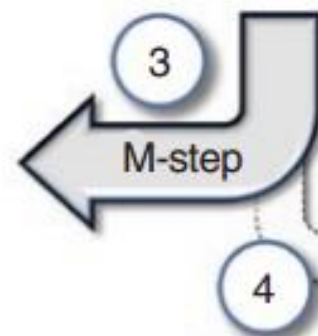
$$(.65 \times .45) / (.65 \times .45 + .55 \times .55)$$

0.45 x		0.55 x	
0.80 x		0.20 x	
0.73 x		0.27 x	
0.35 x		0.65 x	
0.65 x		0.35 x	

Coin A	Coin B
$\approx 2.2 \text{ H}, 2.2 \text{ T}$	$\approx 2.8 \text{ H}, 2.8 \text{ T}$
$\approx 7.2 \text{ H}, 0.8 \text{ T}$	$\approx 1.8 \text{ H}, 0.2 \text{ T}$
$\approx 5.9 \text{ H}, 1.5 \text{ T}$	$\approx 2.1 \text{ H}, 0.5 \text{ T}$
$\approx 1.4 \text{ H}, 2.1 \text{ T}$	$\approx 2.6 \text{ H}, 3.9 \text{ T}$
$\approx 4.5 \text{ H}, 1.9 \text{ T}$	$\approx 2.5 \text{ H}, 1.1 \text{ T}$
$\approx 21.3 \text{ H}, 8.6 \text{ T}$	$\approx 11.7 \text{ H}, 8.4 \text{ T}$

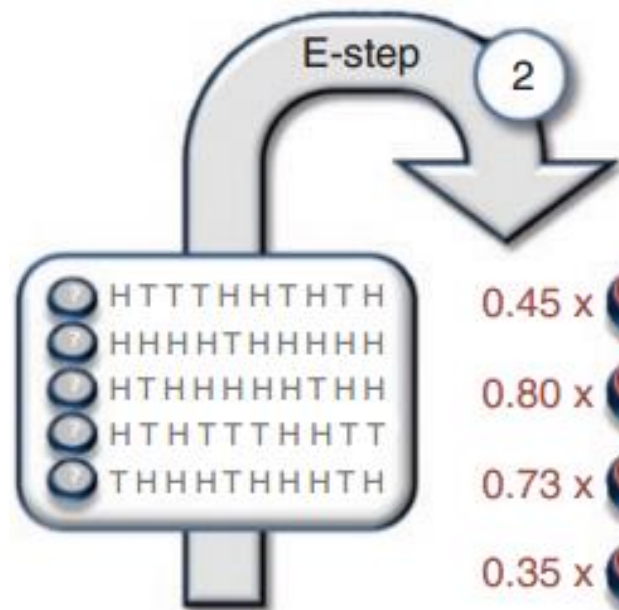
$$\hat{\theta}_A^{(1)} \approx \frac{21.3}{21.3 + 8.6} \approx 0.71$$

$$\hat{\theta}_B^{(1)} \approx \frac{11.7}{11.7 + 8.4} \approx 0.58$$



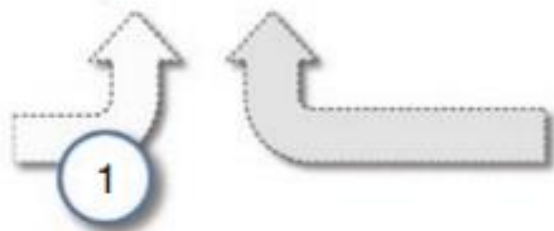
$$\hat{\theta}_A^{(10)} \approx 0.80$$

$$\hat{\theta}_B^{(10)} \approx 0.52$$



$$\hat{\theta}_A^{(0)} = 0.60$$

$$\hat{\theta}_B^{(0)} = 0.50$$



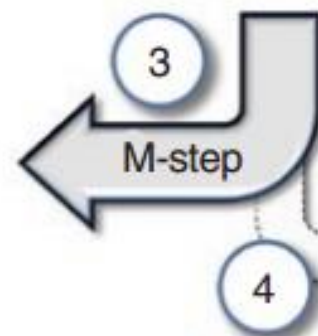
$$(.55 \times .55) / (.65 \times .45 + .55 \times .55)$$

0.45 x	0.55 x
0.80 x	0.20 x
0.73 x	0.27 x
0.35 x	0.65 x
0.65 x	0.35 x

Coin A	Coin B
$\approx 2.2 \text{ H}, 2.2 \text{ T}$	$\approx 2.8 \text{ H}, 2.8 \text{ T}$
$\approx 7.2 \text{ H}, 0.8 \text{ T}$	$\approx 1.8 \text{ H}, 0.2 \text{ T}$
$\approx 5.9 \text{ H}, 1.5 \text{ T}$	$\approx 2.1 \text{ H}, 0.5 \text{ T}$
$\approx 1.4 \text{ H}, 2.1 \text{ T}$	$\approx 2.6 \text{ H}, 3.9 \text{ T}$
$\approx 4.5 \text{ H}, 1.9 \text{ T}$	$\approx 2.5 \text{ H}, 1.1 \text{ T}$
$\approx 21.3 \text{ H}, 8.6 \text{ T}$	$\approx 11.7 \text{ H}, 8.4 \text{ T}$

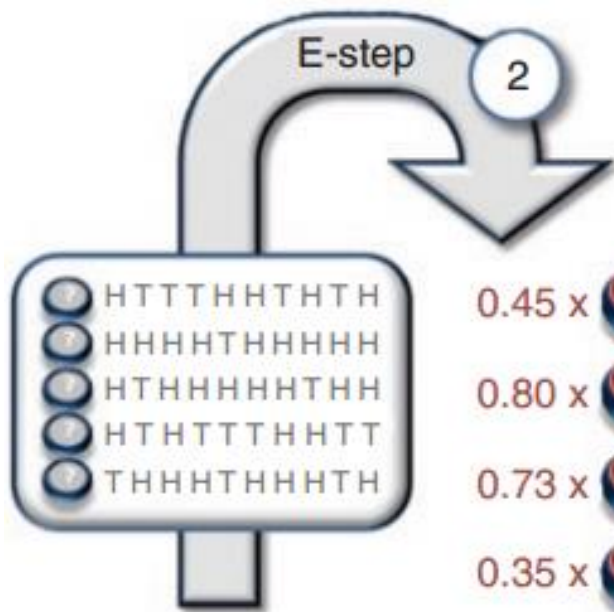
$$\hat{\theta}_A^{(1)} \approx \frac{21.3}{21.3 + 8.6} \approx 0.71$$

$$\hat{\theta}_B^{(1)} \approx \frac{11.7}{11.7 + 8.4} \approx 0.58$$



$$\hat{\theta}_A^{(10)} \approx 0.80$$

$$\hat{\theta}_B^{(10)} \approx 0.52$$



$$\hat{\theta}_A^{(0)} = 0.60$$

$$\hat{\theta}_B^{(0)} = 0.50$$

0.45 x

0.55 x

0.80 x

0.20 x

0.73 x

0.27 x

0.35 x

0.65 x

0.65 x

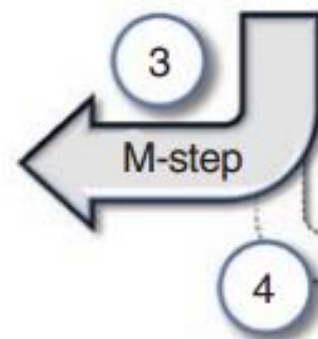
0.35 x

.45(5)

.8(9)

.55(5)

Coin A	Coin B
$\approx 2.2 \text{ H}, 2.2 \text{ T}$	$\approx 2.8 \text{ H}, 2.8 \text{ T}$
$\approx 7.2 \text{ H}, 0.8 \text{ T}$	$\approx 1.8 \text{ H}, 0.2 \text{ T}$
$\approx 5.9 \text{ H}, 1.5 \text{ T}$	$\approx 2.1 \text{ H}, 0.5 \text{ T}$
$\approx 1.4 \text{ H}, 2.1 \text{ T}$	$\approx 2.6 \text{ H}, 3.9 \text{ T}$
$\approx 4.5 \text{ H}, 1.9 \text{ T}$	$\approx 2.5 \text{ H}, 1.1 \text{ T}$
$\approx 21.3 \text{ H}, 8.6 \text{ T}$	$\approx 11.7 \text{ H}, 8.4 \text{ T}$



$$\hat{\theta}_A^{(1)} \approx \frac{21.3}{21.3 + 8.6} \approx 0.71$$

$$\hat{\theta}_B^{(1)} \approx \frac{11.7}{11.7 + 8.4} \approx 0.58$$

.65(4)

$$\hat{\theta}_A^{(10)} \approx 0.80$$

$$\hat{\theta}_B^{(10)} \approx 0.52$$

حداکثر انتظار

1. در گام اول برای یکی از دو دسته مجهول (معمولا پارامترها) مقدار تصادفی در نظر گرفته می شود.
2. مجهول دیگر از روی مقادیر داده شده به دست می آید. (Expectation)
3. از اطلاعات به دست آمده برای مجهول دوم اطلاعات مجهول اولیه بهینه می شود. (Maximization)
4. مراحل ۲ و ۳ به صورت تکرار شونده انجام می شود تا همگرا شود.

حداکثر انتظار (تاریخچه)

حداکثر انتظار به طور پراکنده مورد استفاده بود تا اینکه Dempster در سال ۱۹۷۷ آن را به صورت کلاسه در آورد.

تعداد مقالات منتشر شده در مجلات IEEE که از EM استفاده می کنند از سال ۱۹۸۸ تا ۲۰۰۶ رشدنمایی داشته است.

در سال ۲۰۰۶ هفتاد مقاله مجله در سال

○ کاربردهای حداکثر انتظار

بازیابی و بخش بندی تصویر و ویدئو، مدل سازی تصویر، همزمانی فرکانس های حامل در مخابرات، تشخیص صوت، ژنتیک

محاسبه احتمال (مثال ۳)

- دو تاس که مستقل از هم هستند باهم برای ۱۰۰۰۰۰ بار پرتاب می شوند.
نتیجه هر پرتاب به صورت زوج (x_1, x_2)
معلوم است احتمال ۱ تا ۶ آمدن را برای هر تاس معین نمایید.

(x_1, x_2)	$x_2 = 1$	$x_2 = 2$	$x_2 = 3$	$x_2 = 4$	$x_2 = 5$	$x_2 = 6$
$x_1 = 1$	(1, 1)	(1, 2)	(1, 3)	(1, 4)	(1, 5)	(1, 6)
$x_1 = 2$	(2, 1)	(2, 2)	(2, 3)	(2, 4)	(2, 5)	(2, 6)
$x_1 = 3$	(3, 1)	(3, 2)	(3, 3)	(3, 4)	(3, 5)	(3, 6)
$x_1 = 4$	(4, 1)	(4, 2)	(4, 3)	(4, 4)	(4, 5)	(4, 6)
$x_1 = 5$	(5, 1)	(5, 2)	(5, 3)	(5, 4)	(5, 5)	(5, 6)
$x_1 = 6$	(6, 1)	(6, 2)	(6, 3)	(6, 4)	(6, 5)	(6, 6)
$f(x_1, x_2)$	$x_2 = 1$	$x_2 = 2$	$x_2 = 3$	$x_2 = 4$	$x_2 = 5$	$x_2 = 6$
$x_1 = 1$	3790	3773	1520	1498	2233	2298
$x_1 = 2$	3735	3794	1497	1462	2269	2184
$x_1 = 3$	4903	4956	1969	2035	2883	3010
$x_1 = 4$	2495	2519	1026	1049	1487	1451
$x_1 = 5$	3820	3735	1517	1498	2276	2191
$x_1 = 6$	6369	6290	2600	2510	3685	3673

محاسبه احتمال (مثال ۳) ادامه

- دو تاس که مستقل از هم هستند باهم برای ۱۰۰۰۰۰ بار پرتاب می شوند.
نتیجه هر پرتاب به صورت زوج (x_1, x_2)
معلوم است احتمال ۱ تا ۶ آمدن را برای هر تاس معین نمایید.

$f_1(x_1)$	x_1	$\tilde{p}_1(x_1)$	x_1	$f_2(x_2)$	x_2	$\tilde{p}_2(x_2)$	x_2
15112	1	0.15112	1	25112	1	0.25112	1
14941	2	0.14941	2	25067	2	0.25067	2
19756	3	0.19756	3	10129	3	0.10129	3
10027	4	0.10027	4	10052	4	0.10052	4
15037	5	0.15037	5	14833	5	0.14833	5
25127	6	0.25127	6	14807	6	0.14807	6

حداکثر انتظار (مثال ۴)

○ دو تاس که مستقل از هم هستند باهم

برای ۱۰۰۰۰۰ بار پرتاب می شوند.

جمع نتیجه هر پرتاب به صورت x_1
 $+ x_2$

معلوم است احتمال ۱ تا ۶ آمدن را برای
هر تاس معین نمایید.

$f(y)$	y
3790	2
7508	3
10217	4
10446	5
12003	6
17732	7
13923	8
8595	9
6237	10
5876	11
3673	12

$$\begin{aligned} f(4) &= f(1, 3) + f(2, 2) + f(3, 1) \\ &= 1520 + 3794 + 4903 = \underline{\underline{10217}} \end{aligned}$$

حداکثر انتظار (مثال ۴) انتخاب مقدار تصادفی

○ ابتدا احتمال رخ داد ۱ تا ۶ برای هرتاس به صورت تصادفی در نظر گرفته می شود. (انتخاب مقدار تصادفی برای مجهول)

$p_{01}(x_1)$	x_1
0.18	1
0.19	2
0.16	3
0.13	4
0.17	5
0.17	6

$p_{02}(x_2)$	x_2
0.22	1
0.23	2
0.13	3
0.16	4
0.14	5
0.12	6

حداکثر انتظار (مثال ۴) مرحله انتظار

○ محاسبه احتمال رخ داد دو عدد

$p_0(x_1, x_2)$	$x_2 = 1$	$x_2 = 2$	$x_2 = 3$	$x_2 = 4$	$x_2 = 5$	$x_2 = 6$
$x_1 = 1$	0.0396	0.0414	0.0234	0.0288	0.0252	0.0216
$x_1 = 2$	0.0418	0.0437	0.0247	0.0304	0.0266	0.0228
$x_1 = 3$	0.0352	0.0368	0.0208	0.0256	0.0224	0.0192
$x_1 = 4$	0.0286	0.0299	0.0169	0.0208	0.0182	0.0156
$x_1 = 5$	0.0374	0.0391	0.0221	0.0272	0.0238	0.0204
$x_1 = 6$	0.0374	0.0391	0.0221	0.0272	0.0238	0.0204

$$p_0(1, 3) = p_{01}(1) \cdot p_{02}(3) = 0.18 \cdot 0.13 = \underline{\underline{0.0234}}$$

$$p_0(2, 2) = p_{01}(2) \cdot p_{02}(2) = 0.19 \cdot 0.23 = \underline{\underline{0.0437}}$$

$$p_0(3, 1) = p_{01}(3) \cdot p_{02}(1) = 0.16 \cdot 0.22 = \underline{\underline{0.0352}}$$

حداکثر انتظار (مثال ۴) مرحله انتظار

○ محاسبه احتمال رخ داد جمع دو عدد

$p_0(y)$	y
0.0396	2
0.0832	3
0.1023	4
0.1189	5
0.1437	6
0.1672	7
0.1272	8
0.0867	9
0.0666	10
0.0442	11
0.0204	12

حداکثر انتظار (مثال ۴) مرحله انتظار

○ محاسبه تعداد حالت های همزمانی دو سکه

$f_q(x_1, x_2)$	$x_2 = 1$	$x_2 = 2$	$x_2 = 3$	$x_2 = 4$	$x_2 = 5$	$x_2 = 6$
$x_1 = 1$	3790	3735.95	2337.03	2530.23	2104.91	2290.74
$x_1 = 2$	3772.05	4364.45	2170.03	2539.26	2821	2495.63
$x_1 = 3$	3515.53	3233.08	1737.39	2714.95	2451.85	1903.39
$x_1 = 4$	2512.66	2497.49	1792.29	2276.72	1804.26	1460.92
$x_1 = 5$	3123.95	4146.66	2419.01	2696.47	2228.84	2712
$x_1 = 6$	3966.37	4279.79	2190.88	2547.24	3164	3673

$$f_q(1, 3) = f(4) \cdot \frac{p_0(1, 3)}{p_0(4)} = 10217 \cdot \frac{0.0234}{0.1023} = \underline{\underline{2337.03}}$$

$$f_q(2, 2) = f(4) \cdot \frac{p_0(2, 2)}{p_0(4)} = 10217 \cdot \frac{0.0437}{0.1023} = \underline{\underline{4364.45}}$$

$$f_q(3, 1) = f(4) \cdot \frac{p_0(3, 1)}{p_0(4)} = 10217 \cdot \frac{0.0352}{0.1023} = \underline{\underline{3515.53}}$$

حداکثر انتظار (مثال ۴) مرحله پیشینه سازی

○ محاسبه احتمال رخ داد هر عدد

$f_{q1}(x_1)$	x_1	$f_{q2}(x_2)$	x_2
16788.86	1	20680.56	1
18162.42	2	22257.42	2
15556.19	3	12646.63	3
12344.34	4	15304.87	4
17326.93	5	14574.86	5
19821.28	6	14535.68	6

$$f_{q1}(1) = f_q(1, 1) + f_q(1, 2) + f_q(1, 3) + f_q(1, 4) + f_q(1, 5) + f_q(1, 6)$$

حداکثر انتظار (مثال ۴) مرحله پیشینه سازی

○ محاسبه احتمال رخ داد هر عدد

$p_{11}(x_1)$	x_1
0.167889	1
0.181624	2
0.155562	3
0.123443	4
0.173269	5
0.198213	6

$p_{12}(x_2)$	x_2
0.206806	1
0.222574	2
0.126466	3
0.153049	4
0.145749	5
0.145357	6

حداکثر انتظار (مثال ۴) در انتها

1584th EM iteration.

$p_{1584,1}(x_1)$	x_1	$p_{1584,2}(x_2)$	x_2
0.158396	1	0.239281	1
0.141282	2	0.260559	2
0.204291	3	0.104026	3
0.0785532	4	0.111957	4
0.172207	5	0.134419	5
0.24527	6	0.149758	6

K-MEANS از دیدگاه EM

○ خوشه بندی k-means

○ تعدادی داده داریم:

$$\{x_1, x_2, x_3, \dots, x_n\}$$

○ هدف: خوشه بندی داده ها به K خوشه متمایز به صورتی که تابع زیر کمینه شود.

○
$$y = \sum_{n=1}^N \sum_{k=1}^K r_{nk} \|x_n - \mu_k\|^2$$

○
$$r_{nk} = \begin{cases} 1 & \text{if } x_n \in S_k \\ 0 & \text{if } x_n \notin S_k \end{cases}$$

K-MEANS از دیدگاه EM

○ خوشه بندی k-means

1. انتخاب مقدار اولیه برای μ_k

2. (Expectation)

با ثابت در نظر گرفتن μ_k مقدار r_{nk} محاسبه می شود.

3. (Maximization)

ثابت نگه داشتن دسته ها و پیدا کردن μ_k

$$\mu_k = \frac{\sum_n r_{nk} x_n}{\sum_n r_{nk}}$$

تمرین

سه سکه در اختیار داریم
فعالیت زیر برای ۱۰ بار انجام شده است.
سکه صفر انداخته شد.

اگر رو آمد سکه یک، سه بار پرتاب می شود.

اگر پشت آمد سکه دو، سه بار پرتاب می شود.

در هر بار نتیجه پرتاب سکه صفر معلوم نیست اما نتیجه پرتاب سکه های یک یا دو در اختیار ماست.

مطلوبست تعیین p_0 ، p_1 و p_2 (احتمال رو آمدن برای هر سکه)

<THT>

<HTH>

<HHH>

<HTH>

<HHT>

<HTT>

<TTH>

<THH>

<HHT>

<HHH>

تمرین

سه سکه در اختیار داریم
فعالیت زیر برای ۱۰ بار انجام شده است.
سکه صفر انداخته شد.

اگر رو آمد سکه یک، سه بار پرتاب می شود.

اگر پشت آمد سکه دو، سه بار پرتاب می شود.

در هر بار نتیجه پرتاب سکه صفر معلوم نیست اما نتیجه پرتاب سکه های یک یا دو در اختیار ماست.

مطلوبست تعیین p_1 ، p_2 و p_3 (احتمال رو آمدن برای هر سکه)

<THT>

<HTH>

<HHH>

<HTH>

<HHT>

<HTT>

<TTH>

<THH>

<HHT>

<HHH>

$$\theta_0 = .2, \theta_1 = .3, \theta_2 = .6$$

$$\alpha = \theta_0 \bar{\theta}_1 \theta_1 \bar{\theta}_1 = 0.2 \times .7 \times .3 \times .7 = 0.0294$$

$$\beta = \bar{\theta}_0 \bar{\theta}_2 \theta_2 \bar{\theta}_2 = 0.8 \times .4 \times .6 \times .4 = 0.0768$$

	$P(Z = H X, \theta)$	$P(Z = T X, \theta)$
THT	$\frac{\alpha}{\alpha + \beta} = 0.277$	$\frac{\beta}{\alpha + \beta} = 0.723$
HTH	0.0986	0.9014
HHH	0.0303	.9697
HTH	0.0986	0.9014
HHT	0.0986	0.9014
HTT	0.277	0.723
TTH	0.277	0.723
THH	0.0986	0.9014
HHT	0.0986	0.9014
HHH	0.0303	.9697

	$P(Z = H X, \theta)$	$P(Z = T X, \theta)$
THT	0.277	0.723
HTH	0.0986	0.9014
HHH	0.0303	.9697
HTH	0.0986	0.9014
HHT	0.0986	0.9014
HTT	0.277	0.723
TTH	0.277	0.723
THH	0.0986	0.9014
HHT	0.0986	0.9014
HHH	0.0303	.9697

$$\theta_1 = \frac{1 \times 3 \times 0.277 + 2 \times 5 \times 0.0986 + 3 \times 2 \times 0.0303}{3 \times 3 \times 0.277 + 3 \times 5 \times 0.0986 + 3 \times 2 \times 0.0303} = 0.4812$$

$$\theta_2 = \frac{1 \times 3 \times 0.723 + 2 \times 5 \times 0.9014 + 3 \times 2 \times 0.9697}{3 \times 3 \times 0.723 + 3 \times 5 \times 0.9014 + 3 \times 2 \times 0.9697} = 0.6578$$

$$\theta_0 = \frac{3 \times 0.277 + 5 \times 0.0986 + 2 \times 0.0303}{10} = 0.1385$$