فصل صفر: احتمال

وقوع احتمال: مربوط به زمانی است که

- از نتیجه یک رخداد بی اطلاعی وجود داشته باشد
- موضوع یا نتیجه رخداد برای فرد مهم باشد

تعاریف اولیه:

آزمایش تصادفی: هر عمل یا رخداد که از نتیجه آن بی اطلاع باشیم را یک آزمایش تصادفی گویند.

مثال:

پرتاب سکه

دو مرتبه پرتاب تاس

زمان خراب شدن یک وسیله برقی سالم

تعداد پرتاب سکه تا مشاهده اولین شیر

انتخاب یک عدد تصادفی از بازه $(0,5)$

فضای نمونه: مجموعه است شامل تمام حالات ممکن آزمایش تصادفی (S)

پرتاب سکه $S = \{H, T\}$

دو مرتبه پرتاب تاس $S=\{(1,1),(1,2),\dots,(1,6),(2,1),\dots,(2,6),\dots,(6,6)\}$

زمان خراب شدن یک وسیله برقی سالم $S=(0,\infty)$

تعداد پرتاب سکه تا مشاهده اولین شیر $S=\{H,TH,TTH,TTTH,\dots\}$

انتخاب یک عدد تصادفی از بازه $(0,5)$؟ $S=(0,5)$

نکته: هر فضای نمونه‌ای با احتمال رخداد یکسان برای هر عضو آنرا فضای نمونه همگن گویند. در غیر اینصورت فضای نمونه ناهمگن می باشد.

پیشامد: هر زیرمجموعه اندازه‌پذیر از فضای نمونه‌ای را پیشامد گویند. $A,B,C,\dots$

روابط بین پیشامدها

- ✓ پیشامد A رخ ندهد: A' یا $A'=S-A$
- ✓ پیشامدها A و B رخ دهند: $A\cap B$
- ✓ پیشامدها A یا B رخ دهند: $A\cup B$
- ✓ پیشامدها A رخ دهد و B رخ ندهند: $A-B=A\cap B'$
- ✓ پیشامدها A یا B رخ دهند ولی هر دو رخ ندهند (تفاضل متقارن)

$$A\Delta B=(A\cup B)-(A\cap B)=(A-B)\cup(B-A).$$

میدان[1] و سیگما-میدان[2]

تعریف میدان: هر مجموعه $\mathcal{F}_0$ شامل زیر مجموعه‌های مجموعه مرجع M که دارای خواص زیر باشد را میدان گویند.

- $M \in \mathcal{F}_0$
- برای هر $A \in \mathcal{F}_0$، داشته باشیم $A' \in \mathcal{F}_0$
- هرگاه $A_1, \dots, A_n \in \mathcal{F}_0$، آنگاه $\bigcup_{i=1}^{n} A_i \in \mathcal{F}_0$،

تعریف سیگما- میدان: هر مجموعه $\mathcal{F}$ شامل زیر مجموعه‌های مجموعه مرجع M که دارای خواص زیر باشد را سیگما- میدان گویند.

- $M \in \mathcal{F}$
- برای هر $A \in \mathcal{F}$، داشته باشیم $A' \in \mathcal{F}$
- هرگاه $A_1, A_2, \dots \in \mathcal{F}$، آنگاه $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$،

مثال: برای فضای نمونه‌ای آزمایش پرتاب سکه، میدان‌های زیر را می توان معرفی نمود.

$$\mathcal{F}_{01} = \{\phi, \{H, T\}\}$$

$$\mathcal{F}_{02} = \{\phi, \{H, T\}, \{H\}, \{T\}\}.$$

مثال: برای فضای نمونه‌ای آزمایش پرتاب دو تاس، میدان‌های زیر را می توان معرفی نمود.

---

[1] Field
[2] σ-Field

$$\mathcal{F}_{01} = \left\{\phi, \{(1,1), (1,2), \dots, (1,6), (2,1), \dots, (2,6), \dots, (6,6)\}\right\},$$

$$\mathcal{F}_{02} = \left\{\begin{matrix}\phi, \{(1,1), (1,2), \dots, (1,6), (2,1), \dots, (2,6), \dots, (6,6)\}, \\ \{(1,1), (1,2), \dots, (1,6)\}, \{(2,1), \dots, (2,6), \dots, (6,6)\}\end{matrix}\right\},$$

$$\mathcal{F}_{03} = \left\{\begin{matrix}\phi, \{(1,1), (1,2), \dots, (1,6), (2,1), \dots, (2,6), \dots, (6,6)\}, \\ \{(1,1)\}, \{(1,2), \dots, (1,6), (2,1), \dots, (2,6), \dots, (6,6)\}\end{matrix}\right\},$$

$$\vdots$$

$$\mathcal{F}_{0n} = \left\{\begin{matrix}\phi, \{(1,1)\}, \{(1,2)\}, \dots, \{(6,6)\}, \\ \{(1,1), (1,2)\}, \{(1,1), (1,3)\}, \dots, \{(5,6), (6,6)\}, \\ \{(1,1), (1,2), (1,3)\}, \{(1,1), (1,2), (1,4)\}, \dots, \{(4,6), (5,6), (6,6)\}, \\ \vdots \\ \{(1,1), (1,2), \dots, (1,6), (2,1), \dots, (2,6), \dots, (6,6)\}\end{matrix}\right\}$$

نکته: کوچکترین میدان یک مجموعه همانند M برابر است با $\mathcal{F}_0 = \{\phi, M\}$.

مثال: سیگما-میدان آزمایش تصادفی زمان خراب شدن یک وسیله برقی سالم را مشخص نمایید.

$$\mathcal{F}_{01} = \{\phi, (0, \infty)\},$$

$$\mathcal{F}_{02} = \{\phi, (0,1), (0.5, 2), [1, \infty), (0,1) \cup [2, \infty), (0.5, 1), (0,0.5] \cup [1, \infty), \dots, (0, \infty)\},$$

$$\mathcal{F} = \{\phi, (0,1), (0.5,2), (1,4), \dots, (0, \infty)\}.$$

نکته: زمانیکه مجموع مرجع S متناهی باشد (تعداد اعضاء) بزرگترین میدان دارای $2^{n(S)}$ عضو می باشد.

نکته: سیگما میدان در علم آمار، مجموعه ای است شامل تمام پیشامدهای ممکن و قابل تعریف از یک آزمایش تصادفی. زین پس این مجموعه را با حرف $\mathcal{F}$ نمایش می دهیم (بدون در نظر گرفتن اینکه این مجموعه میدان یا سیگما-میدان واقعی باشد)

انواع احتمال

1- فراوانی: $P(A) = \frac{\#(A)}{\#(Practice)}$

2- تجربی:

3- نظری:

ویژگی‌های تابع احتمال

- $\forall A \subset S, 0 \leq P(A) \leq 1$
- $P(S) = 1$
- $\forall A, B \subset S,\ \ that\ \ A \cap B = \phi\ \ \ then\ \ P(A \cup B) = P(A) + P(B)$

بنابراین در هر آزمایش تصادفی با یک ۳تایی مرتب بصورت $(S, \mathcal{F}, P)$ سروکار خواهیم داشت. به این ۳تایی مرتب **فضای احتمال** گویند. به بیان دیگر هرگاه با ۳تایی به فرم $(S, \mathcal{F}, P)$ مواجه باشیم بطوریکه S مجموعه‌ای شامل تمام حالات ممکن یک آزمایش تصادفی، $\mathcal{F}$ مجموعه‌ای شامل تمام پیشامدهای ممکن از فضای نمونه‌ای S و P که تابع احتمال تعریف شده بر $\mathcal{F}$ می باشد، آنگاه با یک مساله احتمالی مواجه هستیم.

توجه مهم: دامنه تابع احتمال، یک سیگما-میدان می باشد، یعنی $P: \mathcal{F} \to [0, 1]$.

مثال: فرض کنید آزمایشی تصادفی با فضای نمونه‌ای $S=\{a_1,\dots,a_5\}$ باشد، بطوریکه $p(a_i)=\frac{i}{15}$.

الف) احتمال آنرا بیابید که اندیس‌ها کمتر از ۶ باشد؟

A: اندیس‌ها کمتر از ۶

$$A=S\Longrightarrow P(A)=1.$$

ب) پیشامد اینکه اندیس‌ها عدد اول باشد را مشخص نمایید.

$$B=\{a_2,a_3,a_5\},$$
$$P(B)=\frac{10}{15}=\frac{2}{3}.$$

فصل اول: متغیر تصادفی

تعریف: هر تابع از فضای نمونه‌ای یک آزمایش تصادفی به مجموعه اعداد حقیقی را متغیر تصادفی گویند. بطور معمول متغیر تصادفی را با حروف بزرگ X، Y، Z و ... نمایش می دهیم

$$X:S\to R.$$

توجه: متغیر تصادفی جنبه عددی یک آزمایش تصادفی است.

مثال: متغیرهای تصادفی زیر را در نظر بگیرید.

الف) X: تعداد شیر قابل مشاهده در پرتاب ٤ سکه

ب) Y: مجموع اعداد رو شده در پرتاب دو تاس

ج) M: ماکزیمم اعداد رو شده در پرتاب دو تاس

د) Z: تفاضل دو عدد تصادفی انتخابی از بازه (۴, ۸)

تکیه‌گاه متغیر تصادفی: زیر مجموعه‌ای از اعداد حقیقی است که شامل تمامی مقادیر ممکن متغیر تصادفی می باشد.

نکته: تکیه‌گاه را با S_X، S_Y و ... نمایش می دهند

مثال: تکیه‌گاه متغیرهای تصادفی تعریف شده در مثال قبل را مشخص نمایید.

$$S_X = \{0, 1, 2, 3, 4\},$$

$$S_Y = \{2, 3, 4, \dots, 12\},$$

$$S_M = \{1, 2, \dots, 6\},$$

$$S_Z = (-4, 4).$$

انواع متغیر تصادفی:

متغیر تصادفی گسسته: هرگاه تکیه‌گاه یک متغیر تصادفی متناهی یا شمارش‌پذیر باشد، متغیر تصادفی را گسسته گویند.

متغیر تصادفی پیوسته: هرگاه تکیه‌گاه یک متغیر تصادفی بازه یا اجتماعی از بازه‌ها باشد، متغیر تصادفی را پیوسته گویند.

مثال: متغیر تصادفی X، Y و M گسسته و متغیر تصادفی Z پیوسته می باشند.

مثال: آزمایشی تصادفی که احتمال موفقیت در آن ۰٫٤ می باشد را آنقدر تکرار می نماییم تا اینکه حالت موفقیت را مشاهده نماییم. متغیر تصادفی D تعداد این آزمایشات را شمارش می نماید. بنابراین

$$S_D = \{1, 2, 3, 4, \dots\}،$$

بنابراین متغیر تصادفی D نیز گسسته می باشد.

تعریف پیشامد: براساس یک متغیر تصادفی همانند X بصورت

$$\{X = a\} \text{ یا } (X = a)$$

نمایش داده می شوند و به این معنا است که متغیر تصادفی X مقداری برابر a داشته باشد.

پیشامد $(a < X \leq b)$ بدین معنا است که متغیر تصادفی X بیشتر از a و حداکثر برابر b شود.

مثال: $(X \leq 2)$: پیشامد آنکه در پرتاب ٤سکه حداکثر ۲شیر مشاهده شود.

$(Y = 6)$: پیشامد آنکه مجموع اعداد روشده در پرتاب دو تاس برابر ۶شود.

$(M > 2)$: پیشامد آنکه ماکزیمم اعداد مشاهده شده در پرتاب دو تاس بیش از ۲ باشد.

$(D = 2k; \quad k \epsilon N)$: پیشامد آنکه اولین موفقیت در پرتاب‌های زوج مشاهده شود.

$\{Z \in (-1,2)\}$: پیشامد آنکه تفاضل دو عدد تصادفی انتخاب شده از بازه $(4,8)$ متعلق به بازه $(-1,2)$ شود یا به بیان دیگر تفاضل بیش از -۱ و کمتر از ۲ باشد.

مثال: با فرض سالم بودن همه سکه‌ها داریم

$$P(X=0)=\frac{1}{16},\ P(X=1)=\frac{\binom{4}{1}}{2^4}=\frac{4}{16},\dots,\ P(X=4)=\frac{\binom{4}{4}}{16}=\frac{1}{16}.$$

برای متغیرهای تصادی گسسته می توان از جدول توزیع احتمال بهره برد

$X=x$	0	1	2	3	4
$P(X=x)$	1/16	4/16	6/16	4/16	1/16

می توان نتایج مربوط به احتمال متغیر تصادفی X را بصورت یک تابع احتمال نوشت

$$P(X=x)=\frac{\binom{4}{x}}{2^4};\quad x=0,1,\dots,4.$$

جدول توزیع احتمال متغیر تصادفی M بصورت

$M=m$	۱	۲	۳	٤	۵	۶
$P(M=m)$	1/36	3/36	5/36	7/36	9/36	11/36

$$P(M=m)=\frac{2m-1}{6^2};\quad m=1,2,\dots,6.$$

جدول توزیع احتمال متغیر تصادفی D بصورت

$D=d$	۱	۲	۳	۴	...
$P(D=d)$	0.4	$0.6*0.4$	$0.6^2*0.4$	$0.6^3*0.4$	...

$$P(D=d)=0.6^{d-1}\times 0.4;\quad d=1,2,3,....$$

برای متغیر تصادفی Y جدول توزیع احتمال را مشخص نمایید.

$Y=y$	2	3	4	5	6	7	8	...	11	12
$P(Y=y)$	1/36	2/36	3/36	4/36	5/36	6/36	5/36	...	2/36	1/36

تابع احتمال آن بفرم زیر می باشد

$$P(Y=y)=\begin{cases}\dfrac{(y-1)}{36}, & y=1,\dots,6\\ \dfrac{(13-y)}{36}, & y=7,\dots,12\end{cases},$$

مثال: در پرتاب دو تاس متغیر تصادفی N را اختلاف اعداد روشده تعریف می نماییم.

الف) تکیه گاه این متغیر تصادفی را بدست آورید.

$$S_N=\{0,1,\dots,5\}$$

ب) جدول توزیع احتمال متغیر تصادفی N را مشخص نمایید

$N=n$	0	1	2	3	4	5
$P(N=n)$	6/36	10/36	8/36	6/36	4/36	2/36

تابع احتمال متغیر تصادفی N برابر است با

$$P(N=n)=\begin{cases}\frac{1}{6} & n=0\\ \frac{(6-n)}{18} & n=1,\dots,5\end{cases}.$$

نتیجه: در متغیرهای تصادفی گسسته می توان بمنظور نمایش مقادیر احتمال مرتبط از جدول توزیع احتمال یا تابع احتمال بهره برد. به بیان دیگر در متغیر تصادفی گسسته مفهومی تحت عنوان تابع احتمال قابل تعریف می باشد.

مثال: دو فرد مسن هر روز در بین ساعات ۹-۱۰ صبح برای ورزش صبحگاهی وارد یک پارک می شوند. اگر متغیرهای تصادفی X و Y را بعنوان زمان ورود فرد اول و دوم تعریف نماییم. آنگاه این دو متغیر تصادفی از نوع پیوسته می باشند و تکیه‌گاه آنها برابر خواهد بود با

$$S_X=S_Y=(9,10).$$

الف) اگر اختلاف زمانی ورود این دو نفر مد نظر باشد و آنرا با حرف Z نمایش دهیم، آنگاه تکیه گاه Z برابر خواهد بود با

$$Z=|X-Y|,\qquad S_Z=(0,1).$$

ب) اگر ورود این دو نفر مستقل از یکدیگر باشد، احتمال آنرا بیابید که نفر اول قبل از نفر دوم وارد پارک شود؟

$$P(X<Y)+P(Y<X)=1,$$

$$P(X<Y)=P(Y<X)\Longrightarrow P(X<Y)=0.5.$$

ج) اگر هر فرد وارد شده به پارک بمنظور ورزش کردن، بین ۱۵ تا ٤۵ دقیقه در پارک بماند، احتمال آنرا بیابید که نفر اول قبل از ساعت ۱۰ از پارک خارج شود.

T: مدت زمان اقامت افراد در پارک $S_T = (15, 45)$.

$$
\begin{aligned}
P(X+T<10) &= \int_9^{10} P(X=a, T<10-a)da \\
&= \int_9^{9:15} P(X=a)da + \int_{9:15}^{9:45} P(X=a)P(T<10-a)da \\
&+ 0 = \frac{1}{4} + \int_{9:15}^{9:45} P(X=a)P(T<10-a)da.
\end{aligned}
$$

انواع توزیع گسسته

۱- توزیع برنولی

هرگاه آزمایش تصادفی دو حالت داشته باشد به آن آزمایش تصادفی دو-دوئی (برنولی) گویند. حال اگر یکی از حالت‌های آزمایش مدنظر باشد (آن حالت را موفقیت گویند) و متغیر تصادفی که تعداد مطلوب در یک مرتبه انجام آزمایش برنولی را شمارش نماید، متغیر تصادفی برنولی گویند. خواهیم داشت

$$S_X = \{0, 1\}.$$

با فرض آنکه احتمال مطلوب را با p نشان دهیم، داریم

$$P(X=1) = p.$$

$$P(X=0) = 1 - p (= q).$$

می توان تابع احتمال فوق را بصورت زیر نوشت

$$P(X=k)=p^k(1-p)^{1-k}; \qquad k=0,1.$$

بصورت خلاصه می نویسیم $X \sim Ber(p)$.

مثال: در یک جعبه ۲۰ مهره با شماره‌های ۱ تا ۲۰ وجود دارد. از این جعبه یک مهره به تصادف انتخاب می نماییم. اگر متغیر تصادفی X تعداد مهره انتخابی و کمتر از ۷ را شمارش کند، تابع احتمال متغیر تصادفی X را بدست آورید.

X: تعداد مهره انتخابی با مقدار کمتر از ۷ $\qquad S_X=\{0,1\}$

$$p=P(X=1)=\frac{6}{20}=0.3, \quad q=1-p=0.7$$

$$P(X=k)=0.3^k 0.7^{1-k}, \quad k=0,1.$$

ب) هرگاه دو مهره همزمان از جعبه خارج نماییم و متغیر تصادفی ***Y*** را تعداد مهره کمتر از ۷ در نظر بگیریم، تکیه‌گاه و جدول توزیع احتمال متغیر تصادفی را بدست آورید.

$$S_Y=\{0,1,2\},$$

Y=y	0	1	2
P(Y=y)	$\frac{\binom{14}{2}}{\binom{20}{2}}$	$\frac{\binom{6}{1}\binom{14}{1}}{\binom{20}{2}}$	$\frac{\binom{6}{2}}{\binom{20}{2}}$

$$P(Y=y)=\frac{\binom{6}{y}\binom{14}{2-y}}{\binom{20}{2}}, \qquad y=0,1,2.$$

ج) اگر دو مهره مرحله قبل را با جایگذاری انتخاب نماییم، جدول توزیع احتمال را کامل نمایید.

$Y=y$	0	1	2
$P(Y=y)$	0.7^2	2*0.3*0.7	0.3^2

$$P(Y=y)=\binom{2}{y}0.3^y(1-0.3)^{2-y},\qquad y=0,1,2.$$

۲- توزیع دوجمله‌ای

اگر یک آزمایش تصادفی برنولی را مستقلا nمرتبه تکرار نماییم و متغیر تصادفی (Y) مدنظر تعداد موفقیت در این آزمایشات را شمارش نماید، خواهیم داشت

Y: تعداد حالت مطلوب در n آزمایش برنولی مستقل

$$S_Y=\{0,1,2,\dots,n\}.$$

و دارای تابع احتمال

$$P(Y=k)=\binom{n}{k}p^k(1-p)^{n-k};\quad k=0,1,2,\dots,n.$$

بصورت خلاصه داریم، $Y\sim Bin(n,p)$.

مثال: می دانیم ۱۵درصد مردم یک شهر مبتلا به دیابت هستند. از بین آنها به تصادف ۲۰ نفر را انتخاب می کنیم، و متغیر تصادفی Y را تعداد مبتلایان به دیابت در بین این ۲۰ نفر در نظر می گیریم، آنگاه

$$S_Y=\{0,1,2,\dots,20\}.$$

$$P(Y=y)=\binom{20}{y}(0.15)^{y}\ \ (0.85)^{20-y};\qquad y=0,1,2,\ldots 20.$$

$$Y \sim Bin(20, 0.15).$$

ب) فرض کنید برای نفر iام، متغیر تصادفی Y_i را بعنوان تعداد دیابتی در نمونه مدنظر تعریف نماییم.

$$S_{Y_i}=\{0,1\} \Longrightarrow Y_i \sim Ber(0.15)\qquad i=1,2,\ldots,20.$$

تعریف می نماییم $Y_1+Y_2+\cdots+Y_{20}$ که بدین معنا است که تعداد افراد دیابتی در ۲۰ نمونه بدست آمده. بنابراین داریم

$$Y=Y_1+Y_2+\cdots+Y_{20}.$$

بدین معنا است که متغیر تصادفی دو جمله‌ای حاصل مجموعه n متغیر تصادفی برنولی مستقل می باشد.

۳- توزیع هندسی

اگر یک آزمایش تصادفی برنولی را تا مشاهده اولین موفقیت تکرار نماییم، به متغیر تصادفی (Z) که این تعداد آزمایش را شمارش می نماید، متغیر تصادفی هندسی گویند. خواهیم داشت

Z: تعداد آزمایش برنولی مستقل تا مشاهده اولین حالت مطلوب

$$S_Z=\{1,2,3,\ldots\}.$$

که دارای تابع توزیع احتمالی برابر با

$$P(Z=k)=p(1-p)^{k-1};\qquad k=1,2,3,\ldots.$$

بصورت خلاصه می نویسیم $Z \sim G(p)$.

مثال: فرض کنید می خواهیم آنقدر آزمایش انجام دهیم تا به اولین فرد دیابتی بر برخوریم. حال اگر متغیر تصادفی Z تعداد افراد انتخاب شده تا مشاهده اولین فرد دیابتی تعریف شود، آنگاه

$$S_Z = \{1, 2, 3, \dots\}$$

$$P(Z = z) = (0.15)(0.85)^{z-1}; \quad z = 1,2,3,\dots,$$

مثال: بطور متوسط از هر ۱۰۰۰۰ نفر یکنفر مبتلا به بیماری غیر قابل علاج H و ۶۰ نفر مبتلا به بیماری I می باشند. در یک تحقیق میدانی در نظر داریم که افراد را بطور تصادفی انتخاب و مورد آزمایش قرار دهیم.

p_1: احتمال ابتلا به بیماری H ($p_1 = \frac{1}{10000} = 0.0001$)

p_2: احتمال ابتلا به بیماری I ($p_2 = \frac{60}{10000} = 0.006$)

الف) احتمال آنرا بیابید که فرد مبتلا به بیماری H در ۱۰۰امین بیمار به بعد مشاهده شود. به بیان دیگر حداقل ۱۰۰ بیمار تا مشاهده اولین فرد مبتلا به بیماری H نیاز باشد.

Z_1: تعداد انتخاب افراد تا مشاهده یک شخص مبتلا به بیماری H ($Z_1 \sim G(0.0001)$)

Z_2: تعداد انتخاب افراد تا مشاهده یک شخص مبتلا به بیماری I ($Z_2 \sim G(0.006)$)

$$
\begin{aligned}
P(Z_1 \geq 100) &= P(Z_1 = 100) + P(Z_1 = 101) + \cdots \\
&= 0.0001(0.9999)^{99} + 0.0001(0.9999)^{100} \\
&+ 0.0001(0.9999)^{101} + \cdots \\
&= 0.0001(0.9999)^{99} \times \frac{1}{1 - 0.9999} = (1 - p_1)^{99} \\
&= (0.9999)^{99}.
\end{aligned}
$$

نتیجه: برای متغیر تصادفی $Z \sim G(p)$، داریم

$$P(Z \geq k) = (1 - p)^{k-1}.$$

به بیان بهتر

$$F_Z(k) = P(Z \leq k) = 1 - P(Z \geq k + 1) = 1 - p^k.$$

ب) احتمال آنرا بیابید که فرد مبتلا به بیماری I قبل از فرد مبتلا به بیماری H رویت شود.

$$
\begin{aligned}
P(Z_1 > Z_2) = P(Z_2 = 1\ \&Z_1 \ge 2) + P(Z_2 = 2\ \&Z_1 \ge 3) + \cdots \\
&= \sum_{t=1}^{\infty} P(Z_2 = t\ \&Z_1 \ge t+1) \\
&= \sum_{t=1}^{\infty} P(Z_2 = t)P(Z_1 \ge t+1) \\
&= \sum_{t=1}^{\infty} 0.006(0.994)^{t-1}(0.9999)^{t} \\
&= 0.006(0.9999)^1 \sum_{t=1}^{\infty} (0.994)^{t-1}(0.9999)^{t-1} \\
&= 0.006(0.9999)^1 \sum_{t=1}^{\infty} (0.994 \times 0.9999)^{t-1} \\
&= 0.006(0.9999)^1 \sum_{t=1}^{\infty} (0.9939006)^{t-1} \\
&= 0.006(0.9999)^1 \times \frac{1}{1-0.9939006} = 0.98360.
\end{aligned}
$$

ج) احتمال آنرا بیابید که شخص با بیماری I در نمونه با شماره فرد و شخص با بیماری H در شماره‌های زوج مشاهده شود.

$$
\begin{aligned}
P(Z_1 = 2k, Z_2 = 2k-1 \quad k \in \mathcal{N}) \\
&= P(Z_1 = 2k,\ \ k \in \mathcal{N})P(Z_2 = 2k-1,\ \ k \in \mathcal{N}) \\
&= 0.49997 \times 0.50150 \cong 0.25
\end{aligned}
$$

$$
\begin{aligned}
P(Z_1 = 2k,\ \ k \in \mathcal{N}) = \sum_{k=1}^{\infty} P(Z_1 = 2k) \\
&= \sum_{k=1}^{\infty} (0.0001)(0.9999)^{2k-1} = \frac{(0.0001)(0.9999)}{1-0.9999^2} \\
&= \frac{0.9999}{1.9999} \cong 0.49997
\end{aligned}
$$

$$P(Z_2 = 2k-1,\ \ k \in \mathcal{N}) = \sum_{k=1}^{\infty} P(Z_2 = 2k-1)$$
$$= \sum_{k=1}^{\infty} (0.006)(0.994)^{2k-2} = \frac{0.006}{1-0.994^2} = \frac{1}{1.994}$$
$$\cong 0.50150.$$

د) احتمال آنرا بیابید که شخص با بیماری I در نمونه با شماره فرد و شخص با بیماری H در نمونه بعد از شخص I مشاهده شود.

$$P(Z_1 = 2k, Z_2 = 2k-1\ \ \ k \in \mathcal{N})$$
$$= \sum_{k=1}^{\infty} P(Z_1 = 2k) P(\, Z_2 = 2k-1)$$
$$= \sum_{k=1}^{\infty} 0.0001(0.9999)^{2k-1} 0.006(0.994)^{2k-2}$$
$$= 0.0001 \times 0.006(0.9999) \sum_{k=1}^{\infty} (0.9939006)^{2k-2}$$
$$= 0.0001 \times 0.006(0.9999) \times \frac{1}{1-(0.9939006)^2}$$
$$= 0.000049.$$

ه) احتمال آنرا بیابید که این دو بیمار بطور متوالی مشاهده شوند.

$$P(Z_1 = k, Z_2 = k+1 \ \ k \in \mathcal{N}) + P(Z_2 = k, Z_1 = k+1 \ \ k \in \mathcal{N}) =$$
$$= \sum_{k=1}^{\infty} P(Z_1 = k) P(\ Z_2 = k+1)$$
$$+ \sum_{k=1}^{\infty} P(Z_2 = k) P(\ Z_1 = k+1)$$
$$= \sum_{k=1}^{\infty} (0.0001)(0.9999)^{k-1}(0.006)(0.994)^{k}$$
$$+ \sum_{k=1}^{\infty} (0.006)(0.994)^{k-1}(0.0001)(0.9999)^{k}$$
$$= 6 \times 10^{-7}(0.994) \sum_{k=1}^{\infty} (0.9939006)^{k-1} + 6$$
$$\times 10^{-7}(0.9999) \sum_{k=1}^{\infty} (0.9939006)^{k-1}$$
$$= 6 \times 10^{-7}(0.994 + 0.9999) \sum_{k=1}^{\infty} (0.9939006)^{k-1}$$
$$= 6 \times 10^{-7}(1.9939) \frac{1}{1 - 0.9939006} = 0.00019614.$$

توزیع دوجمله‌ای منفی

در نظر داریم آزمایش برنولی را آنقدر تکرار نماییم تا به rامین موفقیت دست یابیم.

حال اگر متغیرتصادفی X تعداد این آزمایشات را شمارش نماید، خواهیم داشت

X: تعداد آزمایش برنولی تا رسیدن به rامین موفقیت.

$$S_X = \{r, r+1, r+2, \dots\}.$$

تابع احتمال این متغیر برابر است با

$$P(X = k) = p \binom{k-1}{r-1} p^{r-1}(1-p)^{k-r}$$

$$= \binom{k-1}{r-1} p^r (1-p)^{k-r}; \quad k = r, r+1, r+2, \dots$$

بصورت خلاصه می توان نوشت: $X \sim NBin(r, p)$.

مثال: در تحقیق مربوط به دیابت در نظر داریم، آنقدر افراد را مورد آزمایش قرار دهیم تا ۱۵امین فرد مبتلا به دیابت شناسایی شود. حال تعریف می کنیم، متغیر تصادفی X را بعنوان تعداد افراد انتخاب شده تا مشاهده ۱۵امین فرد مبتلا به دیابت. می دانیم

$$S_X = \{15, 16, 17, \dots\},$$

ب) احتمال آنرا بیابید که این اتفاق در ٤٨امین فرد نمونه رخ دهد.

$$P(X=48) = (0.15) \times \binom{47}{14} (0.15)^{14} (0.85)^{33} = \binom{47}{14} 0.15^{15} (0.85)^{33},$$

ج) احتمال آنرا بیابید که در این تحقیق بیش از ٤٠ نفر غیر مبتلا مشاهده کنیم.

$$\begin{aligned} P(X > 55) &= P(X=56) + P(X=57) + P(X=58) + \cdots \\ &= \sum_{k=56}^{\infty} \binom{k}{14} (0.15)^{15} (0.85)^{k-15} \\ &= (0.15)^{15} \sum_{k=56}^{\infty} \binom{k}{14} (0.85)^{k-15}. \end{aligned}$$

مثال: اگر موضوع مورد سوال، بررسی امکان مشاهده ۱۰امین فرد مبتلا به بیماری *I* قبل از مشاهده اولین فرد مبتلا به بیماری *H* باشد. احتمال آنرا بدست آورید.

ب) احتمال آنرا بیابید که ۱۰امین فرد مبتلا به بیماری *I* در نمونه ۸۷۰۰ بدست آید، و دومین فرد مبتلا به بیماری *H* در نمونه ۱۵۰۰۰ بدست آید؟

٤- توزیع فوق هندسی

اگر جامعه به دو دسته a و b تایی تقسیم شود و به تصادفی از این جامعه nتا را انتخاب نماییم (بدون جایگذاری) و متغیر تصادفی X تعداد اعضاءِ دسته aتایی انتخاب شده در این نمونه را شمارش می نماید. آنگاه

$$S_X = \{max(0, n-b), \ldots, min(n, a)\}.$$

$$P(X=k) = \frac{\binom{a}{k}\binom{b}{n-k}}{\binom{a+b}{n}}; \qquad k \in S_X.$$

بصورت خلاصه می نویسیم: $X \sim HG(a, b; n)$.

مثال: فرض کنید در یک جمع ٤ (b) فوتبالیست و ۶ نفر (a) علاقمند به والیبال حضور دارند. از بین آنها ۷ نفر را بتصادف انتخاب می کنیم (یک گروه ۷نفره ایجاد می کنیم)، متغیر تصادفی X را تعداد علاقمندان به والیبال انتخاب شده در این گروه تخصیص می دهیم. موارد زیر را پاسخ دهید. یعنی $X \sim HG(6, 4; 7)$.

الف) تکیه گاه متغیر تصادفی X

$$S_X = \{3, 4, 5, 6\}.$$

ب) احتمال آنرا بیابید که تعداد والیبالیست‌ها انتخابی بیشتر باشد

$$P(X \geq 4) = P(X=4) + P(X=5) + P(X=6)$$
$$= \frac{\binom{6}{4}\binom{4}{3}}{\binom{10}{7}} + \frac{\binom{6}{5}\binom{4}{2}}{\binom{10}{7}} + \frac{\binom{6}{6}\binom{4}{1}}{\binom{10}{7}}$$

$$P(X \geq 4) = 1 - P(X = 3) = 1 - \frac{\binom{6}{3}\binom{4}{4}}{\binom{10}{7}}.$$

ج) احتمال آنرا بیابید که ۲ نفر خاص از والیبالیست‌ها حتما انتخاب شوند.

$$\frac{\binom{4}{4}\binom{4}{1} + \binom{4}{3}\binom{4}{2} + \binom{4}{2}\binom{4}{3} + \binom{4}{1}\binom{4}{4}}{\binom{10}{7}} = \frac{\binom{8}{5}}{\binom{10}{7}} = \frac{56}{120}.$$

نکته: این احتمال فقط مربوط به دو نفر خاص می باشد (یعنی ارتباطی به اینکه والیبالیست یا فوتبالیست هستند، ندارد)

نکته: رابطه زیر همیشه برقرار می باشد

$$\sum_{k \in S_X} \frac{\binom{a}{k}\binom{b}{n-k}}{\binom{a+b}{n}} = 1 \Leftrightarrow \frac{\sum_{k \in S_X}\binom{a}{k}\binom{b}{n-k}}{\binom{a+b}{n}} = 1$$

$$\Leftrightarrow \sum_{k \in S_X} \binom{a}{k}\binom{b}{n-k} = \binom{a+b}{n}.$$

مثال: از بین ۲۰ عدد که روی ۲۰ توپ نوشته شده است، به تصادف ۵ عدد انتخاب می نماییم. لطفا در ادامه مطالب خواسته شده را با دو روش با جایگذاری و بدون جایگذاری پاسخ دهید. متغیر تصادفی نشان دهنده میانه اعداد انتخابی باشد.

الف) احتمال آنرا بیابید که عدد میانه برابر ۱۱ شود؟ $P(X = 11)$

روش با جایگذاری $\frac{10^2 9^2}{20^5}$ روش بدون جایگذاری $\frac{\binom{10}{2}\binom{9}{2}}{\binom{20}{5}}$

۶. توزیع پواسن (پواسون)

هرگاه متغیر تصادفی (همانند X) تعداد رخداد در زمان و مکان خاص را شمارش نماید، یک متغیر تصادفی پواسن خواهیم داشت.

$$S_X = \{0, 1, 2, \dots\}$$

$$P(X = k) = \frac{e^{-\lambda}\, \lambda^k}{k!}; \qquad k = 0, 1, 2, \dots.$$

که در آن λ متوسط رخداد می باشد.

$$e^a = 1 + a + \frac{a^2}{2!} + \frac{a^3}{3!} + \cdots = \sum_{r=0}^{\infty} \frac{a^r}{r!}; \qquad a \in \mathcal{R}.$$

به صورت مختصر این توزیع را با $X \sim Pos(\lambda)$ نشان می دهیم.

مثال: بطور متوسط در هر یک ساعت ۷ تلفن به میز منشی رییس یک شرکت نفتی وصل می شود. با توجه به این موضوع و اینکه این منشی در روز از ساعت ۸ صبح تا ۱۶ عصر مشغول به کار است، موارد ادامه را پاسخ دهید

الف) متغیر تصادفی X را تعداد تماس به میز منشی در بین ساعات ۹ تا ۱۰ صبح تعریف می نماییم، احتمال آنرا بیابید که در بین این ساعات، منشی ۵ تماس داشته باشد

X: تعداد تماس تلفنی بین ساعات ۹ الی ۱۰ صبح

$$X \sim Pos(7); \quad P(X = 5) = \frac{e^{-7} 7^5}{5!} = 0.1277.$$

تفسیر: در هر ۱۰۰۰۰ هزار روز کاری منشی، بطور متوسط در ۱۲۷۷ روز آن و در بین ساعات ۹ الی ۱۰ صبح تنها ۵ تماس تلفنی را پاسخگو می باشد.

ب) احتمال آنرا بیابید که در بین ساعات ١٤ تا آخر وقت تماسی نداشته باشد

Y: تعداد تماس در بین ساعات ١٤ الی ۱۶ $Y \sim Pos(14)$

$$P(Y=0)=e^{-14}\frac{14^0}{0!}=e^{-14}\cong 0.$$

ج) تعداد تماس در بین ساعات ۸ الی ۹ صبح بیشتر از ۱۲ تا ۱۳:۳۰ باشد

Z: تعداد تماس در بین ساعات ۸ الی ۹ صبح $Z \sim Pos(7)$

R: تعداد تماس در بین ساعات ۱۲ الی ۱۳:۳۰ صبح $R \sim Pos(10.5)$

$$\begin{aligned} P(Z>R) &= \sum_{t=0}^{\infty}\big(P(R=t)P(Z\geq t+1)\big)\\ &= \sum_{t=0}^{\infty}\left(\frac{e^{-10.5}\ (10.5)^t}{t!}\sum_{r=t+1}^{\infty}\frac{e^{-7}\ 7^r}{r!}\right)\\ &= e^{-17.5}\sum_{t=0}^{\infty}\left(\frac{(10.5)^t}{t!}\sum_{r=t+1}^{\infty}\frac{7^r}{r!}\right). \end{aligned}$$

د) احتمال آنرا بیابید که تعداد تماس در دو ساعت اول با دو ساعت آخر کاری یکسان باشد.

S: تعداد تماس در بین ساعات ۸ الی ۱۰ صبح $S \sim Pos(14)$

K: تعداد تماس در بین ساعات ١٤ الی ۱۶ بعدازظهر $K \sim Pos(14)$

$$P(S=K)=\sum_{t=0}^{\infty}P(K=t)P(S=t)=\sum_{t=0}^{\infty}\left(\frac{e^{-14}\ \ (14)^{t}}{t!}\right)^{2}$$
$$=e^{-28}\sum_{t=0}^{\infty}\frac{(14)^{2t}}{t!\,t!}.$$

مثال: متوسط تعداد حاضرین در یک بانک در ساعات کاری آن بصورت جدول زیر می باشد

ساعات کاری	۸-۱۰	۱۰-۱۲:۳۰	۱۲:۳۰-۱٤	۱٤-۱۵	۱۵-۱۶
متوسط مراجعه کنندگان	۲۰	٤۵	۲۷	۳۰	۱۰

الف) احتمال آنکه در ۳ ساعت اول یک روز کاری بیش از ۵۰ نفر به بانک مراجعه نمایند.

X: تعداد مراجعین در بین ساعات ۸ الی ۱۱ صبح $\quad X\sim Pos(38)$

$$P(X>50)=\sum_{k=51}^{\infty}\frac{e^{-38}\,38^{k}}{k!}=e^{-38}\sum_{k=51}^{\infty}\frac{38^{k}}{k!}.$$

تکلیف: حداقل ۵ گزینه برای این مثال طراحی و با پاسخ آنرا بارگذاری نمایید.

مثال: هنگام تایپ یک شخص حرفه‌ای بطور متوسط ۳ اشتباه در ۵صفحه دارد، یک شخص آماتور ۱۰اشتباه و یک شخص تازه‌کار ۲۰ اشتباه تایپی دارد. اگر در بازار بطور متوسط ٤۰درصد تایپیست‌ها حرفه‌ای، ۳۰درصد آماتور و مابقی تازه‌کار باشند.

الف) احتمال آنرا بیابید که در ۱۰ صفحه تایپ شده، ۱۱ اشتباه تایپی مشاهده شود.

X: نوع تایپیست (حرفه‌ای ۱، آماتور ۲ و تازه کار ۳)

$X = x$	1	2	3
$P(X = x)$	0.4	0.3	0.3

Y: تعداد اشتباه تایپی در ۱۰ صفحه

$$Y|X = 1 \sim Pos(6), \quad Y|X = 2 \sim Pos(20), \quad Y|X = 3 \sim Pos(40)$$

$$\begin{aligned} P(Y = 11) &= P(Y = 11, X = 1) + P(Y = 11, X = 2) \\ &+ P(Y = 11, X = 3) \\ &= P(X = 1)P(Y = 11|X = 1) \\ &+ P(X = 2)P(Y = 11|X = 2) \\ &+ P(X = 3)P(Y = 11|X = \quad) \\ &= 0.4 \times \frac{e^{-6}6^{11}}{11!} + 0.3 \times \frac{e^{-20}6^{11}}{11!} + 0.3 \times \frac{e^{-40}6^{11}}{11!} =. \end{aligned}$$

ب) با مشاهده ۱۱ اشتباه در ۱۰ صفحه بنظر شما کدامیک از سطوح کاری، مسئولیت تایپ را داشته اند؟

$$P(X = 1|Y = 11) = \frac{P(X = 1, Y = 11)}{P(Y = 11)} = \frac{P(X = 1)P(Y = 11|X = 1)}{P(Y = 11)}$$

$$= -$$

$$P(X = 2|Y = 11) = \cdots$$

$$P(X = 3|Y = 11) = \cdots$$

متغیر تصادفی‌های پیوسته

تابع چگالی: هر تابع در ریاضیات که در دو شرط زیر صدق نماید را تابع چگالی گویند.

- برای هر مقدار حقیقی نامنفی شود به بیان دیگر $\forall x \in R \quad f(x) \geq 0$
- انتگرال آن وجود و برابر ۱ شود، یعنی $\int_{-\infty}^{\infty} f(x)dx = 1$.

مثال: کدامیک از توابع زیر چگالی هستند

الف) $f(x) = 8x^3 - 2x \quad\quad 0 < x < 1$

ب) $g(x) = e^{-\frac{|x|}{4}}, \quad\quad x \in R$

ج) $h(x) = 5x^6, \quad\quad 1 < x < 4$

جواب)

الف) شرط دوم برقرار است زیرا $\int_0^1 8x^3 - 2xdx = 1$، این در حالی است که تابع f برای $x < 1/2$ مقدارش منفی می شود.

ب) تابع g بطور واضح شرط اول را دارا می باشد. در حالیکه

$$\int_{-\infty}^{\infty} e^{-\frac{|x|}{4}}dx = \int_{-\infty}^{0} e^{\frac{x}{4}}dx + \int_{0}^{\infty} e^{-\frac{x}{4}}dx = 8.$$

ج) برای بازه $(1, 4)$ تابع h یک تابع همیشه مثبت می باشد، یعنی شرط اول برقرار است. برای شرط دوم داریم

$$\int_1^4 5x^6 dx = \frac{5}{7}(4^7 - 1) = 11702.14286.$$

بنابراین نتیجه می گیریم توابع فعلی هیچکدام هر دو شرط چگالی بودن را ندارند. این در حالی است که دو تابع g و h مقدار انتگرال آنها مثبت شده و با تقسیم طرفین تساوی بر مقدار بدست آمده از انتگرالشان، یا تابع چگالی می توان ساخت

این بدین معنا است که از تابع g می توان یک تابع چگالی بفرم

$$g_X(x) = \frac{1}{8} e^{-\frac{|x|}{4}}, \quad x \in R,$$

ساخت. همچنین از تابع h، تابع چگالی زیر را می توان ایجاد نمود.

$$h_Y(y) = \frac{7}{(4^7 - 1)} y^6, \quad 1 < y < 4.$$

نتیجه: توابع در ریاضیات بر دو قسم هستند

۱- **توابعی که چگالی هستند**، یعنی دو خاصیت یک تابع چگالی را دارا می باشند.

۲- **توابعی که چگالی نیستند**. این توابع خود بر دو نوع دارند

- **چگالی شدنی**: یعنی شرط اول را دارا هستند و تنها مقدار انتگرال آنها برابر یک نمی باشد (مقداری متناهی است) و با تقسیم تابع مدنظر بر مقدار حاصل از انتگرال، به یک چگالی تبدیل می شود.
- **چگالی شدنی نیستند**. این نوع توابع غالبا خاصیت یک را ندارند یا آنکه مقدار انتگرال آنها متناهی نیست.

مثال: مقدار m را در توابع زیر طوری بیابید که تابع حاصل چگالی باشد.

الف) $f(x) = x^2 - mx, \qquad 0 < x < 1$

ب) $h(x) = mx^r e^{-\frac{x}{b}}, \quad x > 0$ جاییکه r و b دو مقدار مثبت معلوم می باشند.

ج) $g(x) = mx^{r-1}(1-x)^{b-1}, \quad 0 < x < 1$ جاییکه r و b دو مقدار مثبت معلوم می باشند.

د) $s(x) = me^{-\frac{x^2}{2}}, \quad x \in R$

جواب: الف)

$$\int_{-\infty}^{\infty} f(x)dx = 1 \Leftrightarrow \int_0^1 (x^2 - mx)dx = \frac{1}{3} - \frac{m}{2} = 1 \Leftrightarrow m = -\frac{4}{3}$$

$$f(x) = x^2 + \frac{4}{3}x, \qquad 0 < x < 1.$$

شرط دوم با قرار دادن $m = -\frac{4}{3}$ برقرار می باشد و تنها کافیست شرط اول چگالی بودن بررسی شود. تابع $f(x) = x^2 + \frac{4}{3}x$ در بازه $(0,1)$ همیشه مثبت می باشد، بنابر این تابع چگالی حاصل از قسمت اول برابر است با

$$f_X(x) = x^2 + \frac{4}{3}x, \qquad 0 < x < 1.$$

ب) برای تابع $h(x)$ شرط اول همیشه برقرار می باشد (به ازاءِ مقادیر مثبت m)، باید مقدار انتگرال بدست آید

$$\int_{-\infty}^{\infty} h(x)dx = 1 \Leftrightarrow \int_{0}^{\infty} mx^r e^{-\frac{x}{b}}dx = m\int_{0}^{\infty} x^r e^{-\frac{x}{b}}dx$$
$$= mb^{r+1}\Gamma(r+1) = 1 \Leftrightarrow m = \frac{1}{b^{r+1}\Gamma(r+1)}.$$

تابع گاما: یکی از پرکاربردترین توابع در ریاضیات تابع گاما می باشد، این تابع بفرم انتگرال تعریف می گردد یعنی

$$\Gamma(\alpha) = \int_{0}^{\infty} x^{\alpha-1} e^{-x} dx,$$

برای هر مقادیر $\alpha > 0$ داریم $\Gamma(\alpha+1) = \alpha\Gamma(\alpha)$. همچنین برای $\alpha \in N$ داریم

$$\Gamma(\alpha) = (\alpha-1)!$$

با توجه به تعریف تابع گاما و تغییر متغیر $\frac{x}{b} = t$، داریم

$$\int_{0}^{\infty} x^r e^{-\frac{x}{b}}dx = \int_{0}^{\infty} (bt)^r e^{-t} bdt = b^{r+1}\int_{0}^{\infty} t^r e^{-t}dt$$
$$= b^{r+1}\Gamma(r+1).$$

این بدان معنا است که تابع چگالی حاصل از تابع اولیه h برابر است با

$$h_X(x) = \frac{x^r e^{-\frac{x}{b}}}{b^{r+1}\Gamma(r+1)}, \quad x > 0.$$

$$\int_{-\infty}^{\infty} h(x)dx = 1 \Leftrightarrow \int_{0}^{\infty} mx^r e^{-\frac{x}{b}}dx = m\int_{0}^{\infty} x^r e^{-\frac{x}{b}}dx$$

ج) به ازاءِ مقادیر بازه $(0,1)$ تابع $g(x)$ مقدار مثبت می پذیر و تنها کافیست مقدار انتگرال انرا بدست آوریم.

$$\int_{-\infty}^{\infty} g(x)dx = 1 \Leftrightarrow \int_0^1 mx^{r-1}(1-x)^{b-1}dx$$
$$= m\int_0^1 x^{r-1}(1-x)^{b-1}dx = m\frac{\Gamma(r)\Gamma(b)}{\Gamma(r+b)} = 1 \Leftrightarrow m$$
$$= \frac{\Gamma(r+b)}{\Gamma(r)\Gamma(b)}.$$

تابع بتا: اولین بار اویلر نشان داد،

$$\int_0^1 x^{r-1}(1-x)^{b-1}dx = \frac{\Gamma(r)\Gamma(b)}{\Gamma(r+b)}$$

و تابعی تحت عنوان تابع بتا بفرم

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

معرفی نمود.

این تابع با مقدار بدست آمده برای m می تواند به یک چگالی تبدیل شود. چگالی حاصل از این تابع، برابر است با

$$g_X(x) = \frac{\Gamma(r+b)}{\Gamma(r)\Gamma(b)}x^{r-1}(1-x)^{b-1}, \quad 0 < x < 1.$$

د) برای تابع $s(x)$ به ازاءِ هر عدد حقیقی مقداری مثبت خواهیم داشت، بنابراین شرط اول چگالی بودن را داراست. بمنظور بررسی شرط دوم می بایست مقدار m بدست آید، بنابراین داریم

$$\int_{-\infty}^{\infty} h(x)dx = 1 \Leftrightarrow m\int_{-\infty}^{\infty} e^{-\frac{x^2}{2}}dx = 2m\int_{0}^{\infty} e^{-\frac{x^2}{2}}dx$$

تغییر متغیر بفرم $u = \frac{x^2}{2}$ $\Leftrightarrow x = \sqrt{2u}$، که داریم $dx = \frac{\sqrt{2}}{2\sqrt{u}}du$

$$2m\int_{0}^{\infty} e^{-\frac{x^2}{2}}dx = \sqrt{2}m\int_{0}^{\infty} u^{-1/2}e^{-u}du = \sqrt{2}m\int_{0}^{\infty} u^{\frac{3}{2}-1}e^{-u}du$$
$$= \sqrt{2}m\Gamma\left(\frac{1}{2}\right) \Leftrightarrow m = \frac{1}{\Gamma\left(\frac{1}{2}\right)\sqrt{2}} = \frac{1}{\sqrt{2\pi}}.$$

با توجه به مقدار متناهی انتگرال فوق، داریم

$$S_X(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}, \qquad x \in R.$$

راه حل دوم: برای محاسبه انتگرال $\int_0^\infty e^{-\frac{x^2}{2}}dx$ بصورت زیر می باشد.

$$I = \int_{0}^{\infty} e^{-\frac{x^2}{2}}dx = \int_{0}^{\infty} e^{-\frac{y^2}{2}}dy \Leftrightarrow I^2 = \int_{0}^{\infty} e^{-\frac{x^2}{2}}dx \times \int_{0}^{\infty} e^{-\frac{y^2}{2}}dy$$
$$= \iint_{0}^{\infty} e^{-\frac{(x^2+y^2)}{2}}dxdy.$$

پیشنهاد این است که تغییر متغیر بفرم $x = \mathrm{r}\cos(\theta)$ و $y = \mathrm{r}\sin(\theta)$، حال داریم

$$dxdy = \begin{vmatrix} \cos(\theta) & -\mathrm{r}\sin(\theta) \\ \sin(\theta) & \mathrm{r}\cos(\theta) \end{vmatrix} = rdrd\theta.$$

بنابراین داریم

$$\int_0^{2\pi}\int_0^{\infty} re^{-\frac{r^2}{2}}dr\,d\theta = \int_0^{2\pi} -e^{-\frac{r^2}{2}}|_0^{\infty}d\theta = \int_0^{2\pi} d\theta = 2\pi \Leftrightarrow I^2 = 2\pi$$
$$\Leftrightarrow m = \sqrt{2\pi}.$$

انواع توزیع‌های پیوسته پرکاربرد

۱- توزیع یکنواخت

انتخاب یک عدد تصادفی از بازه $[a,b]$ را با متغیر تصادفی X نمایش می دهیم، بنابراین خواهیم داشت

$$S_X = [a,b],$$
$$f_X(x) = \frac{1}{b-a}; \qquad a \le x \le b.$$

این توزیع را بصورت خلاصه با $X \sim U(a,b)$ نمایش می دهیم.

مثال: یک نرم افزار کامپیوتری بمنظور تولید اعداد تصادفی از بازه $(0,10)$ یک عدد بتصادف انتخاب می نماید.

الف) احتمال آنرا بیابید که عدد انتخابی حداقل ۳ باشد

X: عدد تصادفی انتخاب شده از بازه $(0,10)$، در نتیجه $X \sim U(0,10)$

$$f_X(x) = \frac{1}{10-0} = 0.1; \qquad 0 < x < 10.$$

$$P(X \geq 3) = \int_3^{10} f_X(x)dx = \int_3^{10} 0.1dx = 0.1(x|_3^{10}) = 0.1(10-3) = 0.7$$

ب) اگر دو عدد تصادفی از این بازه انتخاب شود، احتمال آنرا بیابید که اختلاف بین آنها کمتر از ۲ باشد.

Y: عدد انتخاب شده اول $\quad Y \sim U(0,10)$

Z: عدد انتخاب شده دوم $\quad Z \sim U(0,10)$

$$P(|Y-Z| < 2) = P(-2 < Y-Z < 2) = ???.$$

۲- توزیع نمایی

مدت زمان کارکرد یا طول عمر یک قطعه سالم و ساده را نشان می دهد. یا بطور مثال مدت زمان انجام درخواست یک شخص درون صف، مدت زمانی که بیمار درون بیمارستان بستری است و ...

T: طول عمر قطعه $\quad S_T = (0, \infty)$

$$f_T(t) = \frac{1}{\lambda} e^{-\frac{t}{\lambda}}; \qquad t > 0.$$

جاییکه λ متوسط زمان ماندگاری در صف یا متوسط طول عمر قطعه می باشد.

این توزیع را به اختصار با $T \sim Exp(\lambda)$ نمایش می دهند.

مثال: فرض کنید ۱۰ قطعه داریم که همگی از یک نوع هستند و مستقلا در طراحی یک مدار بکار رفته‌اند. با فرض آنکه میانگین طول عمر هر یک از قطعات برابر ۵۰ ساعت باشد، در هر یک از طراحی‌های زیر مطلب خواسته شده را بدست آورید.

الف) قطعه شماره ۱ بیش 75 واحد زمانی کار کند.

T_i: مدت زمان کارکرد قطعه iام ($i = 1. \dots .10$) که داریم $T_i \sim Exp(50)$

$$P(T_1 > 75) = \int_{75}^{\infty} f_{T_1}(t)dt = \int_{75}^{\infty} \frac{1}{50} e^{-\frac{t}{50}} dt = -e^{-\frac{t}{50}}\Big|_{75}^{\infty} = e^{-\frac{75}{50}} = e^{-1.5}.$$

از تابع فوق می توان نتیجه گرفت که برای هر عدد همانند $c > 0$ بجای ۷۵ می توان نوشت

$$P(T_i > c) = e^{-\frac{c}{50}}.$$

$$P(T_i < c) = 1 - P(T_i \geq c) = 1 - e^{-\frac{c}{50}}.$$

ب) در یک طراحی سری از همه قطعات، سیستم بیش از ۲۰ ساعت کار کند.

T^*: طول عمر ۱۰ قطعه نمایی سری بسته شده، یعنی $T^* = \min(T_i; i = 1, \dots, 10)$

$$
\begin{aligned}
P(T^* > 20) &= P(\min(T_i; i = 1, \dots, 10) > 20) \\
&= P(T_1 > 20, T_2 > 20, \dots, T_{10} > 20) \\
&= P(T_1 > 20) P(T_2 > 20) \dots P(T_{10} > 20) \\
&= \underbrace{e^{-\frac{20}{50}} \times e^{-\frac{20}{50}} \times \dots \times e^{-\frac{20}{50}}}_{\text{10 مرتبه}} = \left(e^{-\frac{20}{50}}\right)^{10} = e^{-4} = 0.0183.
\end{aligned}
$$

تفسیر:

ب) در طراحی موازی حاصل از همه قطعات، سیستم حاصل، حداکثر ~~٤۰~~ ۱۰۰ ساعت کار کند.

T^m: طول عمر ۱۰ قطعه موازی بسته شده یعنی $T^m = \max(T_i; i = 1, \dots, 10)$

$$
\begin{aligned}
P(T^m \le 100) &= P(\max(T_i; i = 1, \dots, 10) \le 100) \\
&= P(T_1 \le 100, T_2 \le 100, \dots, T_{10} \le 100) \\
&= P(T_1 \le 100) P(T_2 \le 100) \dots P(T_{10} \le 100) \\
&= \underbrace{\left(1 - e^{-\frac{100}{50}}\right)\left(1 - e^{-\frac{100}{50}}\right) \dots \left(1 - e^{-\frac{100}{50}}\right)}_{\text{10 مرتبه}} = (1 - e^{-2})^{10} \\
&= 0.234.
\end{aligned}
$$

تفسیر: از هر ۱۰۰۰ دستگاهی که حاصل از موازی بستن ۱۰ عدد از این قطعات است، بطور متوسط ۲۳٤ عدد از آنها کمتر از ۱۰۰ ساعت کار می کنند.

ج) دو قسمت ۵ قطعه‌ای سری را موازی در سیستم قرار می دهیم، احتمال آنرا بیابید که بین ۳۰ تا ۱۲۰ ساعت کار کند

Y_i: طول عمر قطعه سری بسته شده‌ی حاصل از ۵ قطعه $i = 1, 2$

Y: طول عمر قطعه

می دانیم $Y = \max(Y_1, Y_2)$

$$P(30 < Y < 120) = P(Y < 120) - P(Y < 30) \\ = (1 - e^{-12})^2 - (1 - e^{-3})^2.$$

$$P(Y < 120) = P(\max(Y_1, Y_2) < 120) = P(Y_1 < 120, Y_2 < 120) \\ = P(Y_1 < 120)P(Y_2 < 120) = (1 - e^{-12})(1 - e^{-12}) \\ = (1 - e^{-12})^2.$$

$$P(Y_1 < 120) = P(\min(T_1. \dots . T_5) < 120) \\ = 1 - P(\min(T_1. \dots . T_5) > 120) \\ = 1 - P(T_1 > 120) \dots P(T_5 > 120) = 1 - \left(e^{-\frac{120}{50}}\right)^5 \\ = 1 - e^{-12}.$$

توزیع پارتو:

هرگاه بخواهیم به مبحث طول عمر یک محصول پس از حداقل زمان حفظ کیفیت بپردازیم، می بایست از توزیع پارتو بهره ببریم. بعنوان مثال شیر پاستوریزه حداقل زمان حفظ کیفیت یا ماندگاری ٤٨ ساعت می باشد. بعد از این مدت باید حتما قبل از مصرف از سالم بودن شیر اطمینان حاصل نمود. این نوع محصولات غالبا در شرایط عادی قابل نگهداری هستند و روی بسته آنها تاریخ تولید و انقضاءِ ثبت می شود.

X: مدت طول عمر یا حفظ کیفیت محصول پس از انقضاءِ زمان مصرف

چگالی این متغیر دارای یک حداقل زمان ماندگاری همانند β می باشد و بفرم زیر است

$$f_X(x) = \frac{\alpha\beta^\alpha}{x^{\alpha+1}}, \qquad x > \beta.$$

مثال: فرض کنید محصول بسته بندی شده لوبیا حداقل زمان حفظ کیفیت در آن ۶ماه باشد. حال اگر فرض کنیم پارامتر $\alpha = 4$ باشد، احتمال آنرا بیابید که این محصول پس از ۲۰۰ روز هنوز دارای کیفیت قابل قبول باشد.

Y: طول عمر محصول لوبیا پس از زمان انقضاء

از مفروضات مثال می توان فهمید $\alpha = 4$ و $\beta = 180$، بنابراین تابع چگالی این متغیر بفرم

$$g_Y(y) = 4\frac{180^4}{y^5}, \qquad y > 180.$$

می باشد. صورت مثال از ما احتمال زیر را خواسته است.

$$P(Y > 200) = \int_{200}^{\infty} g_Y(y)dy = 4\int_{200}^{\infty} \frac{180^4}{y^5}dy = \left(\frac{180}{200}\right)^4 = 0.6561.$$

نتیجه: برای متغیر تصادفی پارتو، X، با پارامترهای α و β خواهیم داشت

$$P(X > t) = \left(\frac{\beta}{t}\right)^\alpha, \qquad P(X \le t) = 1 - \left(\frac{\beta}{t}\right)^\alpha; \qquad t > \beta.$$

ب) با تغییر مقدار α از ٤ به ۱۰ احتمال خواسته شده قبل را مجدد محاسبه نمایید.

$$P(Y>200)=\left(\frac{180}{200}\right)^{10}=0.3487.$$

نتیجه: با افزایش مقدار α، احتمال ماندگاری محصول پس از زمان انقضاءِ کاهش می یابد. یعنی هرچه مقدار این پارامتر بیشتر شود، زمان فسادپذیری یا افت کیفیت محصول بیشتر خواهد بود.

متغیر تصادفی پارتو را بصورت خلاصه با $X\sim pareto(\alpha,\beta)$ نمایش می دهند.

مثال: برای ۵ نمونه از تولیدات شرکتی با حداقل طول عمر ۱۵۰ ساعت و ضریب فسادپذیری ۳، $Z_1,\dots,Z_5$، احتمالات زیر را بدست آورید.

الف) احتمال آنکه زمان فسادپذیری همگی زیر ۲۰۰ ساعت باشد یا به بیان دیگر، طولانیترین زمان حفظ کیفیت در بین این ۵ محصول زیر ۲۰۰ ساعت باشد.

$$\begin{aligned}P\big(Z_{(5)}=\max(Z_1,\dots,Z_5)\le 200\big)&=P(Z_1\le 200,\dots,Z_5\le 200)\\&=P(Z_1\le 200)\dots P(Z_5\le 200)=\left(1-\left(\frac{150}{200}\right)^3\right)^5\\&=0.0646.\end{aligned}$$

تفسیر: بطور متوسط از هر ۱۰۰۰۰ بسته ۵تایی از این محصولات، انتظار داریم ۶۴۶ عدد ازآنها تا قبل از ۲۰۰ ساعت اولیه تولید فاسد شوند.

ب) این ۵ محصول دارای حداقل زمان حفظ کیفیت ۱۸۰ ساعت وحداکثر زمان حفظ کیفیت آنها ۲۲۰ ساعت باشند.

$$
\begin{aligned}
&P(\min(Z_1,\dots,Z_5) > 180, \max(Z_1,\dots,Z_5) < 220)\\
&\qquad\qquad = P(180 < Z_1 < 220, \dots, 180 < Z_5 < 220)\\
&\qquad\qquad = P(180 < Z_1 < 220) \dots P(180 < Z_5 < 220)\\
&\qquad\qquad = (0.261742)^5 = 0.001229.
\end{aligned}
$$

تفسیر: بطور متوسط از هر ۱۰۰۰۰۰۰ بسته ۵تایی از این محصولات انتظار داریم ۱۲۲۹ در بین ۱۸۰ تا ۲۲۰ ساعت اولیه بعد از تولید بطور کامل فاسد شده باشند.

می دانیم

$$
\begin{aligned}
P(180 < Z_1 < 220) &= P(Z_1 < 220) - P(Z_1 < 180)\\
&= \left(1 - \left(\frac{150}{220}\right)^3\right) - \left(1 - \left(\frac{150}{180}\right)^3\right) = \left(\frac{150}{180}\right)^3 - \left(\frac{150}{220}\right)^3\\
&= 0.261742.
\end{aligned}
$$

در توزیع پارتو داریم

$$
P(a < X < b) = \left(\frac{\beta}{a}\right)^\alpha - \left(\frac{\beta}{b}\right)^\alpha.
$$

امید ریاضی[3](متوسط مقدار متغیر تصادفی): اگر از یک متغیر تصادفی بیشمار مشاهده داشته باشیم، میانگین مقدار حاصل از این مشاهدات را امید ریاضی گویند (تعریف مفهومی).

نکته: برای متغیر تصادفی X امیدریاضی آن را بصورت $E(X)$ نمایش می دهیم.

نکته: در حل مسائل مربوط به امیدریاضی می بایست فرمول محاسباتی آن، ابتدائا ذکر شود.

فرمول محاسباتی امید ریاضی در متغیرهای گسسته

$$E(X)=\sum_{x\epsilon S_X} x\,P(X=x).$$

مثال: امیدریاضی متغیرهای تصادفی تعریف شده تا کنون را بدست آورید

$$\begin{aligned}E(X)&=\sum_{x=0}^{4} xP(X=x)\\&=0\times\frac{1}{16}+1\times\frac{4}{16}+2\times\frac{6}{16}+3\times\frac{4}{16}+4\times\frac{1}{16}=2.\end{aligned}$$

تفسیر: در آزمایش تصادفی پرتاب ٤سکه سالم، انتظار داریم بطور متوسط ۲شیر مشاهده نماییم.

[3] Expectation of random variable

$$E(Y)=\sum_{y=2}^{12} yP(Y=y)$$
$$=2\times\frac{1}{36}+3\times\frac{2}{36}+4\times\frac{3}{36}+\cdots+11\times\frac{2}{36}+12\times\frac{1}{36}$$
$$=7.$$

$$E(D)=\sum_{d=1}^{\infty} d\,P(D=d)$$
$$=1\times 0.4+2\times 0.6\times 0.4+3\times 0.4\times 0.6^2+4\times 0.4$$
$$\times 0.6^3+\cdots$$
$$=(0.4+0.6\times 0.4+0.4\times 0.6^2+0.4\times 0.6^3+\cdots)$$
$$+(0.6\times 0.4+0.4\times 0.6^2+0.4\times 0.6^3+\cdots)$$
$$+(0.4\times 0.6^2+0.4\times 0.6^3+\cdots)+(0.4\times 0.6^3+\cdots)+\cdots$$
$$=\frac{0.4}{1-0.6}+\frac{0.6\times 0.4}{1-0.6}+\frac{0.6^2\times 0.4}{1-0.6}+\frac{0.6^3\times 0.4}{1-0.6}+\cdots$$
$$=1+0.6+0.6^2+0.6^3+\cdots=\frac{1}{1-0.6}=\frac{10}{4}=2.5$$

تفسیر: برای رسیدن به اولین موفقیت، بطور متوسط نیازمند انجام ۲,۵ آزمایش هستیم.

$$E\big(g(X)\big)=\sum_{x\epsilon S_X} g(x)P(X=x).$$

مثال: متوسط توان دوم متغیر تصادفی X، متوسط مجذور متغیر تصادفی Y و متوسط معکوس متغیر تصادفی M را بدست آورید.

$$E(X^2)=\sum_{x=0}^{4} x^2P(X=x)=0^2\times\frac{1}{16}+1^2\times\frac{4}{16}+2^2\times\frac{6}{16}+3^2\times\frac{4}{16}+4^2\times\frac{1}{16}=\frac{80}{16}=5.$$

$$E(\sqrt{Y}) = \sum_{y=2}^{12} \sqrt{y}\, P(Y = y) = \sqrt{2}\frac{1}{36} + \sqrt{3}\frac{2}{36} + \cdots + \sqrt{12}\frac{1}{36} = 2.602.$$

$$E\left(\frac{1}{M}\right) = \sum_{m=1}^{6} \frac{1}{m} P(M = m) = \frac{1}{1}\frac{1}{36} + \frac{1}{2}\frac{3}{36} + \frac{1}{3}\frac{5}{36} + \frac{1}{4}\frac{7}{36} + \frac{1}{5}\frac{9}{36} + \frac{1}{6}\frac{11}{36} = 0.265.$$

خواص امید ریاضی

۱- برای هر عدد ثابت همانند a داریم $E(a) = a$

۲- $E(X \pm Y) = E(X) \pm E(Y)$ (امیدریاضی جمع برابر جمع امیدریاضی‌ها است)

۳- $E(aX) = aE(X)$

٤- اگر متغیر تصادفی X نامنفی باشد (یعنی مقادیر تکیه‌گاه آن همیشه مثبت باشند) آنگاه $E(X) \geq 0$.

واریانس: متوسط پراکندگی مقادیر متغیر تصادفی نسبت به امیدریاضی را واریانس گویند. فرمول آن بفرم زیر می باشد

$$Var(X) = E\big(X - E(X)\big)^2 = \sum_{x \epsilon S_X} \big(x - E(X)\big)^2 P(X = x)$$
$$= E(X^2) - \big(E(X)\big)^2 = E(X^2) - E^2(X).$$

زیرا

$$\begin{aligned} Var(X) &= E\big(X - E(X)\big)^2 = E\big(X^2 + E^2(X) - 2XE(X)\big) \\ &= E(X^2) + E\big(E^2(X)\big) - 2E\big(XE(X)\big) \\ &= E(X^2) + E^2(X) - 2E(X)E(X) = E(X^2) - E^2(X). \end{aligned}$$

مثال: واریانس متغیرهای تصادفی X و M را بدست آورید

$$Var(X) = E(X^2) - E^2(X) = 5 - 2^2 = 1.$$

$$Var(M) = E(M^2) - E^2(M) = 21.972 - (4.472)^2 = 1.973.$$

زیرا

$$E(M^2) = \sum_{m=1}^{6} m^2 P(M = m) = 1^2 \frac{1}{36} + 2^2 \frac{3}{36} + \cdots + 6^2 \frac{11}{36} = 21.972.$$

مثال: ضریب چولگی پیرسون متغیر تصادفی M را بدست آورید

(نکته: فرمول چولگی بفرم مقابل می باشد $\rho_M = \frac{E\left(M-E(M)\right)^3}{\left(Var(M)\right)^{\frac{3}{2}}}$).

$$\begin{aligned} E\left(M - E(M)\right)^3 &= \sum_{m=1}^{6} \left(m - E(M)\right)^3 P(M = m) \\ &= \left(1 - E(M)\right)^3 \frac{1}{36} + \left(2 - E(M)\right)^3 \frac{3}{36} + \left(3 - E(M)\right)^3 \frac{5}{36} \\ &+ \cdots + \left(1 - E(M)\right)^6 \frac{11}{36} = \cdots. \end{aligned}$$

خواص واریانس

۱- $Var(aX) = a^2 Var(X)$

$$\begin{aligned} Var(aX) &= E((aX)^2) - \left(E(aX)\right)^2 = E(a^2X^2) - \left(aE(X)\right)^2 \\ &= a^2E(X^2) - a^2\left(E(X)\right)^2 = a^2 \left(E(X^2) - \left(E(X)\right)^2\right) \\ &= a^2 Var(X). \end{aligned}$$

۲- برای هر عدد ثابت همانند a داریم $Var(a) = 0$

$$Var(a) = E(a^2) - \big(E(a)\big)^2 = a^2 - a^2 = 0.$$

۳- $Var(X) \geq 0$

می دانیم $Var(X) = E\big(X - E(X)\big)^2$ و همچنین $\big(X - E(X)\big)^2$ یک عبارت غیر منفی می باشد، بنابر خاصیت ۳ امیدریاضی می بایست واریانس همیشه مقدار مثبت بپزیرد.

٤- $Var(X \pm Y) = Var(X) + Var(Y) \pm 2Cov(X, Y)$

مثال: اگر دو تاس را بریزیم و متغیرهای تصادفی X و Y را به ترییب ماکزیمم و اختلاف اعداد روشده تعریف نماییم. خواهیم داشت

$$S_X = \{1, 2, \dots, 6\}.$$

$$S_Y = \{0, 1, \dots, 5\}.$$

$$P(Y = 1) = \frac{10}{36}. \quad P(X = 4, Y = 2) = \frac{2}{36}. \quad P(X = 6, Y = 3) = \frac{2}{36}.$$

می توان احتمال رخ دادن دو متغیر تصادفی X و Y را در جدول زیر خلاصه نمود

Y \ X	0	1	2	3	4	5
1	$\frac{1}{36}$	0	0	0	0	0
2	$\frac{1}{36}$	$\frac{2}{36}$	0	0	0	0
3	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	0	0	0
4	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	0	0
5	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	0
6	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$

به این جدول، جدول توزیع توام گویند.

Y \ X	0	1	2	3	4	5	$P(X=x)$
1	$\frac{1}{36}$	0	0	0	0	0	$1/36$
2	$\frac{1}{36}$	$\frac{2}{36}$	0	0	0	0	$3/36$
3	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	0	0	0	$5/36$
4	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	0	0	$7/36$

5	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	0	$9/36$
6	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$\frac{2}{36}$	$11/36$
$P(Y=y)$	$6/36$	$10/36$	$8/36$	$6/36$	$4/36$	$2/36$	1

بنابر نتایج در سطر و ستون اضافه این جدول توزیع توام می توان گفت، از جدول توزیع توام متغیرها می توان توزیع هریک از متغیرها را نیز بدست آورد که به این مقادیر توزیع احتمال حاشیه ای گویند.

مثال: فرض کنید جدول توزیع توام دو متغیر تصادفی M و N بفرم زیر باشد

$N \backslash M$	-2	0	2	4	$P(N=n)$
-3	0.10	0.05	0.00	0.05	0.20
-2	0.07	0.03	0.20	0.00	0.30
0	0.00	0.15	0.05	0.02	0.22
1	0.03	0.05	0.10	0.10	0.28
$P(M=m)$	0.20	0.28	0.35	0.17	1

$$S_M = \{-2, 0, 2, 4\}; \quad S_N = \{-3, -2, 0, 1\},$$

$$P(M=0, N=-2) = 0.03.$$

$$P(N=0) = 0.22; \quad P(M=4) = 0.17.$$

$$P(M^2=4) = P(M=-2) + P(M=2) = 0.20 + 0.35 = 0.55.$$

الف) امید ریاضی M (متوسط مقدار متغیر تصادفی M) را بدست آورید.

$$E(M) = \sum_{m \epsilon S_M} m\, P(M = m) = (-2)0.20 + 0 \times 0.28 + 2 \times 0.35 + 4 \times 0.17 = 0.98.$$

ب) احتمال آنرا بیابید که حاصلضرب این دو متغیر تصادفی بیشتر از صفر شود

$$\begin{aligned} P(M \times N > 0) &= P(M = -2, N = -3) + P(M = -2, N = -2) \\ &+ P(M = 2, N = 1) + P(M = 4, N = 1) \\ &= 0.1 + 0.07 + 0.1 + 0.1 = 0.37. \end{aligned}$$

ج) واریانس N را محاسبه نمایید.

$$E(N^2) = (-3)^2 * 0.2 + (-2)^2 * 0.3 + 0^2 * 0.22 + 1^2 * 0.28 = 3.28.$$

$$E(N) = (-3) * 0.2 + (-2) * 0.3 + 0 * 0.22 + 1 * 0.28 = -0.92.$$

$$Var(N) = E(N^2) - E^2(N) = 3.28 - (-0.92)^2 = 2.4336.$$

د) مقدار مربوط به $E(MN)$ را بدست آورید.

$$\begin{aligned} E(MN) = \sum_{m \epsilon S_M} \sum_{n \epsilon S_N} m \times n \times P(M = m, N = n) \\ = (-3)(-2)0.10 + (-3)(0)0.05 + (-3)(2)0.00 + (-3)4 \\ * 0.05 + \cdots + 1 * 4 * 0.10 = 0.02. \end{aligned}$$

M / N	-2	0	2	4
-3	0.10	0.05	0.00	0.05
-2	0.07	0.03	0.2	0.00

0	*0.00*	*0.15*	*0.05*	*0.02*
1	*0.03*	*0.05*	*0.10*	*0.10*

$$E(M^2N) = \sum_{m\epsilon S_M} \sum_{n\epsilon S_N} m^2 \times n \times P(M=m, N=n)$$
$$= (-3)(-2)^2 0.10 + (-3)(0)^2 0.05 + (-3)(2)^2 0.00$$
$$+ (-3)4^2 * 0.05 + \cdots + 1 * 4^2 * 0.10 = -16.8.$$

کوواریانس: میزان تغییرپذیری یک متغیر نسبت به متغیر دیگر را نشان می دهد.

$$Cov(X,Y) = E(XY) - E(X)E(Y).$$

این فرمول نشان میدهد، بمنظور محاسبه کوواریانس می بایست توزیع توام دو متغیر تصادفی مورد مطالعه را داشته باشیم.

ادامه مثال بالا: قسمت ه) کوواریانس دو متغیر تصادفی M و N را محاسبه کنید

$$Cov(M,N) = E(MN) - E(M)E(N)$$
$$= 0.02 - (-0.92)(0.98) = -0.9216.$$

تکالیف جلسه آینده

۱- کوواریانس مربوط به متغیرهای تصادفی X و Y در پرتاب دو تاس را محاسبه نمایید.

۲- برای جدول توزیع توام احتمال زیر موارد خواسته شده را بدست آورید

Z / R	-2	-0.5	1	5	$P(R=r)$
-2	0.02	0.04	0.05	0.00	0.11
-1	0.08	0.00	0.05		0.22
0	0.00	0.03	0.05	0.06	0.14
1	0.02	0.08	0.05		0.25
2	0.03	0.20	0.05	0.00	0.28
$P(Z=z)$	0.15	0.35	0.25	0.25	1

الف) جدول فوق را کامل کنید

ب) امید و واریانس هر یک از متغیرها را بدست آورید.

ج) جدول توزیع احتمال متغیر $X = Z + R$ را ایجاد نمایید.

د) امید و واریانس متغیر تصادفی X را محاسبه کنید.

۳. صورت فرمول ضریب چولگی پیرسون را بسط داده و برای متغیر تصادفی M آنرا محاسبه نمایید.

$$\begin{aligned}
E\big(M-E(M)\big)^3 &= E\big(M^3-3M^2E(M)+3ME^2(M)-E^3(M)\big)\\
&= E(M^3)-3E\big(M^2E(M)\big)+3E\big(ME^2(M)\big)-E\big(E^3(M)\big)\\
&= E(M^3)-3E(M)E(M^2)+3E^2(M)E(M)-E^3(M)\\
&= E(M^3)-3E(M)E(M^2)+2E^3(M).
\end{aligned}$$

ویژگی‌های کوواریانس: برای دو متغیر تصادفی X و Y و همچنین اعداد حقیقی a و b داریم.

۱) $Co\ \ (X,Y) = Cov(Y,X)$

۲) $Cov(X,aY) = aCov(X,Y)$ (کوواریانس به واحد اندازه‌گیری متغیرهای تصادفی وابسته می باشد)

۳) $Cov(a,Y) = 0$

٤) $Cov(X \pm b,Y) = Cov(X,Y)$ (کوواریانس به جابجایی متغیرهای تصادفی وابسته نمی باشد)

۵) $Cov(X+Z,Y) = Cov(X,Y) + Cov(Z,Y)$

۶) $Cov(X,X) = Var(X)$ (واریانس میزان تغییر پذیری یک متغیر نسبت به خودش را نشان می دهد یا به بیان دیگر واریانس نیز یک نوع کوواریانس می باشد) زیرا

$$Cov(X,X) = E(XX) - E(X)E(X) = E(X^2) - E^2(X) = Var(X).$$

نتیجه: $$Var(X \pm Y) = Var(X) + Var(Y) \pm 2Cov(X,Y)$$

$$\begin{aligned} Var(X-Y) &= \mathrm{Cov}(X-Y,X-Y) \\ &= \mathrm{Cov}(X,X-Y) - \mathrm{Cov}(Y,X-Y) \\ &= \mathrm{Cov}(X-Y,X) - \mathrm{Cov}(X-Y,Y) \\ &= \mathrm{Cov}(X,X) - \mathrm{Cov}(Y,X) - \mathrm{Cov}(X,Y) + \mathrm{Cov}(Y,Y) \\ &= \mathrm{Var(X)} + \mathrm{Var(Y)} - 2\mathrm{Cov}(X,Y). \end{aligned}$$

ضریب همبستگی پیرسون

میزان همبستگی دو متغیر تصادفی را نشان می دهد. مقدار آن عددی است در بازه $[-1,1]$ می باشد. بفرم زیر محاسبه می شود

$$\rho_{X.Y} = \frac{Cov(X,Y)}{\sqrt{Var(X)\times Var(Y)}},$$

نحوه تفسیر مقدار حاصل از ضریب همبستگی

مقادیر مثبت نشان دهنده میزان همبستگی دو متغیر تصادفی در یک راستا می باشند، به این معنا که افزایش یکی باعث افزایش دیگری می شود. و مقادیر منفی عکس این مطلب برای آنها صدق می کند.

هرچقدر قدرمطلق این عدد به مقدار یک نزدیک باشد نشاندهنده قدرت یا همبستگی شدیدتر این دو متغیر می باشد. مقادیر نزدیک صفر نشان از عدم همبستگی این دو متغیر دارند.

تکلیف بعدی:

محاسبه ضریب همبستگی برای ۳ جدول توزیع توام متغیرهای تصادفی تعریف می شود.

متغیرهای تصادفی پیوسته

تعریف: هر گاه تکیه‌گاه متغیر تصادفی بازه یا اجتماعی از بازه‌ها باشد، متغیر تصادفی را پیوسته گویند.

نکته: در متغیرهای تصادفی پیوسته احتمال در تک نقطه بی معنا است (تعریف نشده است). به بیان دیگر در متغیرهای تصادفی پیوسته تابع احتمال وجود ندارد.

تابع چگالی: هر تابع همانند $f(x)$ که دارای ویژگی‌های زیر باشد را چگالی گویند

- ✓ به ازاءِ همه مقادیر حقیقی مقدارش غیر منفی باشند: $\forall x \in \mathcal{R}, \quad f(x) \geq 0$
- ✓ $\int_{-\infty}^{\infty} f(x)dx = 1$

مثال: کدامیک از توابع زیر چگالی هستند

الف) $f(x) = \frac{1}{x^2}, \qquad x \geq 1$

$$\int_{-\infty}^{\infty} f(x)dx = \int_{1}^{\infty} \frac{1}{x^2}dx = -\frac{1}{x}|_{1}^{\infty} = 0 - (-1) = 1.$$

بدلیل داشتن هر دو شرط فوق، چگالی است یعنی $f_X(x) = \frac{1}{x^2}$ برای مقادیر $x \geq 1$ یک تابع چگالی است. بنابراین می نویسیم

$$f_X(x) = \frac{1}{x^2} \qquad x \geq 1.$$

ب) $g(x) = 5x^3; \qquad 0 < x < 2$

$$\int_{-\infty}^{\infty} g(x)dx = \int_{0}^{2} 5x^3 dx = \frac{5}{4}x^4|_{0}^{2} = 20.$$

بدلیل نداشتن شرط دوم، چگالی نیست، ولی با تقسیم مقدار مثبت حاصل از انتگرال به طرفین تساوی خواهیم داشت

$$g_X(x) = \frac{x^3}{4}; \qquad 0 < x < 2،$$

که یک چگالی است و به این نوع توابع چگالی شدنی گویند.

ج) $h(x) = 2x - 4x^2 = 2x(1-2x); \quad 0 < x < 1$ بدلیل نداشتن شرط یک چگالی نیست

مثال: در هر یک از موارد زیر مقدار m (در صورت امکان) را طوری بیابید که توابع چگالی شوند.

الف) $f(x) = \frac{m}{x^2}; \qquad x \geq 0$ مقداری برای m قابل بدست آمدن نیست زیرا مقدار انتگرال به ازای هر m بینهایت خواهد شد.

ب) $g(y) = 4y^2 - my; \quad 0 < y < 1$. برای بدست آوردن شرط دوم، انتگرال زیر را می بایست حل بنماییم

$$\int_{-\infty}^{\infty} g(y)dy = \int_0^1 (4y^2 - my)dy = \left(\frac{4}{3}y^3 - \frac{m}{2}y^2\right)\Big|_0^1 = \frac{4}{3} - \frac{m}{2} = 1 \rightarrow m = \frac{2}{3}.$$

مقدار $m=2/3$ شد ولی به ازای این مقدار، تابع $g(y)$ مقادیر منفی در بازه $0 < y < 1$ خواهد پذیرفت (یعنی شرط اول برقرار نیست.). بنابراین مطالب تابع $g(y)$ چگالی شدنی نیست.

ج) $s(z) = mze^{-3z^2}; \quad z > 0$.

مقدار مناسب برای آنکه انتگرال تابع $s(z)$ برابر یک شود، برابر است با $m=6$ و تابع $s_Z(z) = 6ze^{-3z^2}; \quad z > 0$ یک چگالی است.

$\int_{-\infty}^{\infty} s(z)dz = \int_0^{\infty} mze^{-3z^2}dz = m\int_0^{\infty} ze^{-3z^2}dz = -\frac{m}{6}e^{-3z^2}|_0^{\infty} = \frac{m}{6} \Longrightarrow m = 6.$

بنابراین تابع چگالی حاصل برابر خواهد بود با $s_Z(z) = 6ze^{-3z^2}; \quad z > 0$.

متغیر تصادفی توام پیوسته (چگالی توام)

همانطور که در متغیرهای تصادفی گسسته تابع احتمال توام یا جدول توزیع توام مورد بررسی قرار گرفت در ادامه به نحوه مطالعه رفتار دو متغیر تصادفی پیوسته می پردازیم. بدین منظور می بایست یک تابع چگالی توام برای این دو متغیر تعریف نمود.

مثال: کدامیک از توابع توام زیر چگالی است

الف) $f(x,y) = mxy + xy^2; \qquad 0 < x, y < 1$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)dx\,dy = \int_0^1\int_0^1 (mxy + xy^2)dx\,dy$$
$$= \int_0^1 \left(\frac{mx^2y}{2} + \frac{x^2y^2}{2}\right)_0^1 dy = \int_0^1 \left(\frac{my}{2} + \frac{y^2}{2}\right)dy$$
$$= \left(\frac{my^2}{4} + \frac{y^3}{6}\right)_0^1 = \frac{m}{4} + \frac{1}{6}.$$

we want to have $\frac{m}{4} + \frac{1}{6} = 1 \rightarrow m = \frac{20}{6}$

در نتیجه به ازاءِ مقدار بدست آمده برای m تابع فوق یک چگالی می تواند باشد، بنابراین آنرا بفرم

$$f_{X,Y}(x,y) = \frac{20}{6}xy + xy^2; \qquad 0 < x, y < 1.$$

نمایش می دهیم.

ب) $g(x,y) = mx^2y^3; \qquad 0 < x < y < 1$

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} g(x,y)dx\,dy = \int_0^1\int_x^1 mx^2y^3dy\,\mathrm{dx} = \int_0^1\int_0^{\mathrm{y}} mx^2y^3dx\,\mathrm{d}y$$
$$= \int_0^1\left(\frac{mx^3y^3}{3}\right)\Big|_0^{\mathrm{y}}\mathrm{dy} = \int_0^1\frac{my^6}{3}\mathrm{dy} = \left(\frac{\mathrm{m}y^7}{21}\right)_0^1 = \frac{\mathrm{m}}{21} \to \frac{\mathrm{m}}{21}$$
$$= 1 \to \mathrm{m} = 21.$$

تابع $g(x,y)$ به ازاءِ مقدار m برابر با ۲۱ دارای ویژگی‌های توابع چگالی توام می باشد، بنابراین آنرا بفرم

$$g_{X,Y}(x,y) = 21x^2y^3; \qquad 0 < x < y < 1.$$

نمایش می دهیم.

مطلب) چگونه چگالی حاشیه‌ای هر متغیر را از چگالی توام استخراج کنیم؟ پاسخ: کافی است برای بدست آوردن تابع چگالی هر متغیر (چگالی حاشیه‌ای) روی تمام مقادیر ممکن متغیرهای دیگر جمع (انتگرال) می بندیم.

به بیان دیگر، اگر تابع چگالی دو متغیر تصادفی برابر $f_{X,Y}(x,y)$ باشد، بمنظور بدست آوردن تابع چگالی حاشیه‌ای هر متغیر تصادفی، همانند X، بر روی مقادیر متغیر دیگر، Y، می بایست جمع انجام دهیم.

مثال: برای هر یک از چگالی‌های فوق، چگالی حاشیه‌ای را بدست آورید.

الف) $f_{X,Y}(x,y) = \frac{20}{6}xy + xy^2; \qquad 0 < x, y < 1$

چگالی حاشیه‌ای متغیر تصادفی X با جمع چگالی توام روی تمام مقادیر ممکن متغیر Y بدست می آید، یعنی

$$f_X(x) = \int_0^1 \left(\frac{20}{6}xy + xy^2\right)dy = \left(\frac{20}{12}xy^2 + \frac{xy^3}{3}\right)_0^1 = \frac{5}{3}x + \frac{x}{3} = 2x; \quad 0 < x < 1.$$

چگالی حاشیه‌ای متغیر تصادفی Y با جمع چگالی توام روی تمام مقادیر ممکن متغیر X بدست می آید، یعنی

$$f_Y(y) = \int_0^1 \left(\frac{20}{6}xy + xy^2\right)dx = \left(\frac{20}{12}x^2y + \frac{x^2y^2}{2}\right)_0^1 = \frac{5}{3}y + \frac{y^2}{2}; \qquad 0 < y < 1.$$

ب) برای چگالی توام $g_{X,Y}(x,y) = 21x^2y^3; \quad 0 < x < y < 1$، تابع چگالی حاشیه-ای متغیر تصادفی X برابر خواهد بود با

$$g_X(x) = \int_x^1 (21x^2y^3)dy = \left(\frac{21x^2y^4}{4}\right)_x^1 = \frac{21x^2}{4} - \frac{21x^6}{4}; \quad 0 < x < 1.$$

همچنین تابع چگالی حاشیه‌ای متغیر تصادفی Y بصورت زیر بدست خواهد آمد

$$g_Y(y) = \int_0^y (21x^2y^3)dx = (7x^3y^3)_0^y = 7y^6; \qquad 0 < y < 1.$$

امیدریاضی در متغیرهای تصادفی پیوسته

$$E(X)=\int_{-\infty}^{\infty} x f_X(x)dx.$$

این مطلب برای هر تابع مشخصی از متغیر تصادفی X نیز برقرار هست، یعنی

$$E\big(h(X)\big)=\int_{-\infty}^{\infty} h(x) f_X(x)dx.$$

مثال: برای ٤ متغیر تصادفی فوق امیدریاضی را بدست آورید

الف) $E(X)=\int_{-\infty}^{\infty} x f_X(x)dx=\int_0^1 x(2x)dx=\int_0^1 2x^2dx=\left(\frac{2x^3}{3}\right)_0^1=\frac{2}{3}$

$$E(Y)=\int_{-\infty}^{\infty} y f_Y(y)dy=\int_0^1 y\left(\frac{5}{3}y+\frac{y^2}{2}\right)dy=\int_0^1\left(\frac{5}{3}y^2+\frac{y^3}{2}\right)dy=$$
$$\left(\frac{5}{9}y^3+\frac{y^4}{8}\right)_0^1=\frac{5}{9}+\frac{1}{8}=\frac{49}{72}.$$

ب) (انجام این محاسبات بر عهده دانشجو می باشد)

ج) واریانس زیر را بدست آورید.

$$Var(X^2)=E(X^4)-E^2(X^2)=\frac{1}{3}-\left(\frac{1}{2}\right)^2=\frac{1}{12}.$$

$$E(X^4)=\int_{-\infty}^{\infty} x^4 f_X(x)dx=\int_0^1 x^4(2x)dx=2\int_0^1 x^5dx=\left(\frac{2x^6}{6}\right)_0^1=\frac{1}{3}.$$

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f_X(x)dx = \int_0^1 x^2(2x)dx = 2\int_0^1 x^3 dx = \left(\frac{2x^4}{4}\right)_0^1 = \frac{1}{2}.$$

مثال: مقدار ضریب همبستگی دو متغیر تصادفی X و Y در مثال‌های قبل را بدست آورید

الف) می دانیم فرمول ضریب همبستگی برابر است با $\rho_{X,Y} = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$، بنابر این خواهیم داشت

$$\rho_{X,Y} = \frac{-0.0000037}{\sqrt{\frac{1}{18}\times 0.054}} \cong 0.$$

$$Cov(X,Y) = E(XY) - E(X)E(Y) = 0.4537 - \left(\frac{2}{3}\right)\left(\frac{49}{72}\right) = -0.0000037.$$

$$E(XY) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} xy f_{X,Y}(x,y)dx\,dy = \int_0^1\int_0^1 xy\left(\frac{20}{6}xy + xy^2\right)dx\,dy = \int_0^1\int_0^1\left(\frac{20}{6}x^2y^2 + x^2y^3\right)dx\,dy = \int_0^1\left(\frac{20x^3y^2}{18} + \frac{x^3y^3}{3}\right)_0^1 dy = \int_0^1\left(\frac{20y^2}{18} + \frac{y^3}{3}\right)dy = \left(\frac{20y^3}{54} + \frac{y^4}{12}\right)_0^1 = \frac{20}{54} + \frac{1}{12} = 0.4537.$$

$$Var(X) = E(X^2) - \left(E(X)\right)^2 = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{1}{2} - \frac{4}{9} = \frac{1}{18}.$$

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f_X(x)dx = \int_0^1 x^2(2x)\,dx = \int_0^1 2x^3 dx = \left(\frac{2}{4}x^4\right)_0^1 = \frac{1}{2}.$$

$$Var(Y) = E(Y^2) - \left(E(Y)\right)^2 = \frac{62}{120} - \left(\frac{49}{72}\right)^2 = 0.054.$$

$$E(Y^2) = \int_{-\infty}^{\infty} y^2 f_Y(y)dy = \int_0^1 y^2 \left(\frac{5}{3}y + \frac{y^2}{2}\right) dy = \int_0^1 \left(\frac{5}{3}y^3 + \frac{y^4}{2}\right) dy =$$

$$\left(\frac{5}{12}y^4 + \frac{y^5}{10}\right)_0^1 = \frac{5}{12} + \frac{1}{10} = \frac{62}{120}.$$

ب) (انجام این محاسبات بر عهده دانشجو می باشد)