فصل سوم

آمار در شبیه‌سازی

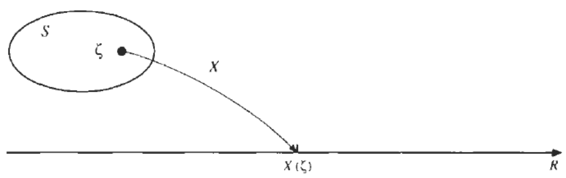
مفاهیم و تعاریف
توزیع های آماری گسسته و پیوسته و مقادیر تصادفی
ساخت اعداد تصادفی
تحلیل داده های ورودی

مفاهیم تعاریف

متغیر تصادفی

متغیر تصادفی تابعی حقیقی است از فضای نمونه به مجموعهٔ اعداد حقیقی که به هر پیشامد فضای نمونه عددی حقیقی نسبت می دهد.

$$X(\varsigma) : S \to R$$



انواع متغیر تصادفی

- متغیر تصادفی گسسته
  - X را متغیر تصادفی گسسته می‌نامند، اگر مقادیری که X می‌گیرد متناهی یا نامتناهی شمارا باشد.
    - تعداد سفارش‌هایی که به کارگاه می‌رسد
    - انداختن یک تاس و آمدن یک عدد خاص
    - تاس ناسالم که احتمال آمدن هر وجه آن با عدد هر وجه متناسب است.

$$0 \le p(x_i) \le 1$$

$$\sum_{x_i} p(x_i) = 1$$

- متغیر تصادفی پیوسته
  - X را متغیر تصادفی پیوسته می‌نامند، اگر مقادیری که X می‌گیرد فاصله‌ای از مجموعه فواصل باشد.
    - عمر یک لامپ

$$f(x) \ge 0$$

$$\int_x f(x) d_x = 1$$

$$\int_{x_0}^{x_0} f(x) d_x = 0$$

تابع توزیع تجمعی

$$F(x)=p(X<x)=\begin{cases}\sum\limits_{x_i<x} p(x_i)\\ \int\limits_{-\infty}^{x} f(t)dt\end{cases}$$

$$\lim_{x\to\infty} F(x)=1$$

$$\lim_{x\to-\infty} F(x)=0$$

$$p(a<x\le b)=F(b)-F(a)\quad ;\quad a<b$$

If X was Continuse stochastic variable then

$$p(a\le x\le b)=p(a\le x\le b)=p(a\le x<b)=p(a<x<b)=F(b)-F(a)=\int_a^b f(x)dx$$

گشتاور، امید ریاضی، واریانس

$$\mu_{(n)}=E(x^n)=\begin{cases}\sum\limits_{\forall i} x_i^{\,n}\, p(x_i)\\ \int\limits_{-\infty}^{+\infty} x^n f(x)dx\end{cases}$$

$$E(x)=\begin{cases}\sum\limits_{x_i} x_i\, p(x_i)\\ \int\limits_{-\infty}^{+\infty} xf(x)dx\end{cases}=\mu$$

$$\mu'_{(n)}=E\left((x-\mu)^n\right)=\begin{cases}\sum\limits_{\forall i}(x-\mu)^n p(x_i)\\ \int\limits_{-\infty}^{+\infty}(x-\mu)^n f(x)dx\end{cases}$$

$$\operatorname{var}(x)=\sigma^2=E\left[(x-E(x))^2\right]=E\left[x^2+E^2(x)-2xE(x)\right]$$

$$E(x^2)+E\left[E^2(x)\right]-2E\left[xE(x)\right]=E(x^2)+E^2(x)-2E^2(x)=E(x^2)-E^2(x)$$

$$=\mu_{(2)}-\mu^2$$

مثال

۱- تاس غیر منصف

X_i	1	2	3	4	5	6
$P(x_i)$	1/21	2/21	3/21	4/21	5/21	6/21
F(X)	1/21	3/21	6/21	10/21	15/21	1

$$E(x)=1\times\frac{1}{21}+2\times\frac{2}{21}+.....+6\times\frac{6}{21}=4.33$$

$$\mathrm{var}(x)=E(x^2)-[E(x)]^2=E(x^2)-4.33^2=21-18.78=2.22$$

۲- عمر لامپ

$$f(x)=\begin{cases}\frac{1}{2}e^{-x/2} x\geq 0\\ o\end{cases}$$

$$p(2\leq x\leq 3)=0.145$$

$$F(x)=\frac{1}{2}\int_{-\infty}^{x}e^{-\frac{t}{2}}dt=\frac{1}{2}\int_{0}^{x}e^{-\frac{t}{2}}dt=1-e^{\frac{-x}{2}}$$

$$E(x)=\frac{1}{2}\int_{0}^{\infty}xe^{\frac{-x}{2}}=2$$

$$Var(x)\Rightarrow\left\{E(x^2)=\frac{1}{2}\int_{0}^{\infty}x^2e^{-\frac{x}{2}}dx=8\right\}\Rightarrow Var(x)=8-2^2=4\rightarrow\sigma=\sqrt{4}=2$$

مد و میانه

مد

- مد در متغیر گسسته مقداری از متغیر تصادفی است که بیشتر از همه روی می‌دهد.
- مد در متغیر پیوسته مقدار ماکسیمم تابع توزیع است

- **میانه**
  - میانه در متغیر تصادفی پیوسته مقداری از متغیر تصادفی است که: F(X<x)=1/2
  - میانه در متغیر تصادفی گسسته اولین xی است که: $p(X<x)\geq\frac{1}{2}$

مثال

- تاس غیر منصف

$$p(X < x) = 15/21 \geq 1/2 \rightarrow x = 5$$

- عمر لامپ

$$\int_0 \frac{1}{2} e^{-\frac{t}{2}} dt = \frac{1}{2}$$

$$\Rightarrow \int_0 \frac{1}{2} e^{-\frac{t}{2}} dt = \left(1 - e^{\frac{-x}{2}}\right) = \frac{1}{2} \Rightarrow e^{\frac{-x}{2}} = \frac{1}{2} \Rightarrow -\frac{x}{2} = \ln\frac{1}{2} \Rightarrow x = 1.3865$$

ضریب چولگی و کشیدگی

با وجود اینکه تعداد گشتاورها نامحدود می باشند ولی در عمل فقط تعداد کمی از آن ها مورد توجه قرار می گیرند.

ضریب مربع

$$C^2[X] = \frac{Var[X]}{E[X]}$$

این معیار نیز برای اندازه گیری میزان پراکندگی یا تغییرات X به کار می رود با این تفاوت که به صورت مقداری بی واحد نرمال شده است.

چولگی

$$v[X] = \frac{E[(X - E[X])^3]}{\sigma^3[X]}$$

اگر توزیع چوله به چپ باشد چولگی منفی و اگر چوله به راست باشد چولگی مثبت و در صورت متقارن بودن چولگی صفر می شود.

کشیدگی
انتهای توزیع

$$k[X] = \frac{E[(X - E[X])^4]}{\sigma^4[X]} - 3$$

اگر کشیدگی انتهای تابع کم باشد این معیار منفی و اگر کشیدگی متوسط باشد صفر و در صورت زیاد بودن کشیدگی، مثبت می باشد.

توزیع های رایج متغیرهای تصادفی گسسته و ساخت مقادیر شبه تصادفی

توزیع یکنواخت گسسته

$$p(x)=\frac{1}{k} \quad ;x=1,2,3,\ldots,k$$

$$E(x)=\sum_{x=1}^{k} x\frac{1}{k}=\frac{1}{k}\sum_{x=1}^{k} x=\frac{1}{k}*\frac{k(k+1)}{2}=\frac{k+1}{2}$$

$$Var(x)=E(x^2)-[E(x)]^2=\sum_{x=1}^{k} x^2\frac{1}{k}-(\frac{k+1}{2})^2$$

$$=\frac{1}{k}*\frac{k(k+1)(2k+1)}{6}-\frac{(k+1)^2}{4}$$

$$=\frac{k^2-1}{12}$$

روش های ساخت مقادیر تصادفی

- روش تبدیل معکوس
- روش تبدیل مستقیم
- روش رد و قبول
- روش پیچش

چگونگی ساخت عدد تصادفی با خصوصیت توزیع یکنواخت گسسته

توزیع برنولی

- متغیر تصادفی برنولی (X) دارای دو نتیجه پیروزی و شکست می باشد. بنابراین فضای نمونه را می توان به شکل S={0,1} در نظر گرفت که در آن ۰ نشانگر شکست و ۱ نشان دهنده پیروزی است. توزیع برنولی به صورت Ber(p) نشان داده می شود که در آن p احتمال موفقیت و در نتیجه 1-p احتمال شکست می باشد. نکات زیر در مورد این متغیر تصادفی قابل استخراج است.

$$p(x)=\begin{cases} p & x=1 \\ 1-p & x=0 \end{cases} \Rightarrow p(x)=p^x q^{1-x} \quad ; \quad x=0,1$$

$$E(x)=1\times p+0\times q=p$$

$$Var(x)=E(x^2)-[E(x)]^2=(1^2\times p+0^2\times q)-p^2=p(1-p)=pq$$

چگونگی ساخت عدد تصادفی با خصوصیت توزیع برنولی

توزیع بینم (دو جمله ای)

فرض کنید n متغیر تصادفی برنولی با هم جمع شوند. حاصل متغیر تصادفی است که می توان آن را با عنوان تعداد پیروزی ها در n آزمایش برنولی تعبیر نمود. تابع توزیع این متغیر تصادفی را می توان به صورت زیر بدست آورد:

$$p(x)=\begin{cases}\binom{n}{x}p^{x}q^{n-x} & x=0,\dots n\\ 0 & \text{otherwise}\end{cases}$$

توزیع بینم (دو جمله ای) ادامه

با توجه به تعریف متغیر تصادفی دو جمله ای داریم:

$$X=X_1+X_2+\dots X_n$$

$$\Rightarrow\begin{cases}E(x)=E\left(\sum_{i=1}^{n}x_i\right)=np\\ \operatorname{var}(x)=\operatorname{var}\left(\sum_{i=1}^{n}x_i\right)=n\operatorname{var}(x_i)=npq\end{cases}$$

$$E(x)=\sum_{x=0}^{n}x\binom{n}{x}p^xq^{n-x}=\sum_{x=0}^{n}n\binom{n-1}{x-1}p^xq^{n-x}=n\sum_{x=0}^{n}\binom{n-1}{x-1}p^{x-1}q^{n-x}\times p=np\sum_{x=0}^{n}\binom{n-1}{x-1}p^{x-1}q^{n-x}=np$$

نکته $\sum_{x=0}^{n}\binom{n}{x}p^xq^{n-x}=(p+q)^n=1^n=1$

$$E[x(x-1)]=\sum_{x=0}^{n}x(x-1)\binom{n}{x}p^xq^{n-x}=n\sum_{x=0}^{n}(x-1)\binom{n-1}{x-1}p^xq^{n-x}=n(n-1)\sum_{x=0}^{n}\left[\binom{n-2}{x-2}p^xq^{n-x}\right]$$

$$=n(n-1)p^2=n^2p^2-np^2$$

$$Var(x)=E[x(x-1)]+E(x)-[E(x)]^2=n^2p^2-np^2+np-n^2p^2=np(1-p)=npq$$

چگونگی ساخت عدد تصادفی با خصوصیت توزیع دو جمله ای

مثال

در یک فرایندساخت، چیپ های نیمه رسانایی با ۲٪ درست معیوب تولید می‌شوند. در این سیستم تولیدی هر روز یک نمونه ۵۰تایی گرفته شده و اگر در نمونه بیشتر از ۲ معیوب باشد فرایند متوقف می‌شود. احتمال توقف فرایند را در هر روز بیابید.

حل: ابتدا بایستی متغیر تصادفی در این سوال تعریف شود.

X تعداد واحدهای ناقص

$$p(x)=\binom{50}{x}(0.02)^x(0.98)^{n-x}$$

$$p(x>2)=1-p(X\le 2)=1-p(x=0)-p(x=1)p(x=2)=1-0.92=0.08$$

$$E(x)=50\times 0.02$$

$$\operatorname{var}(x)=50\times 0.02\times 0.98=0.98$$

توزیع هندسی

آزمایش های برنولی مستقل از هم را در نظر بگیرید. متغیر تصادفی هندسی (X) تعداد آزمایش های برنولی تا رسیدن به اولین موفقیت می باشد. این توزیع به شکل Ge(p) نشان داده می شود. (p احتمال موفقیت و1-p احتمال شکست می باشد). از آنجا که تعداد آزمایش ها نامحدود می باشد فضای حالت به شکل {...,S={1,2,...,k است. تابع احتمال متغیر تصادفی x به شکل زیر می باشد و داریم:

- $P(x)=q^{x-1}p;\ x=1, 2, \ldots$
  - $E(x)=1/p$
  - $\operatorname{var}(x)=q/p^2$

توزیع هندسی به طور گسترده در مدل های ریاضی به علت خاصیت بی حافظگی این توزیع استفاده می شود. بی حافظگی یعنی: $p(x>k+n \mid x>k)=p(x>n) \quad ; \quad n\ge 1$

چگونگی ساخت عدد تصادفی با خصوصیت توزیع هندسی

مثال

در مثال قبل احتمال اینکه سومین نمونه، اولین معیوب باشد را بیابید.

پیروزی: یافتن معیوب

$P(x=3) = 0.98^2 * 0.02$

Prepared By Dr. Kazemipoor

توزیع پواسون

$$p(x) = \binom{n}{x} p^x q^{n-x}$$

$$n \to \infty, \quad np = \lambda$$

$$p(x) = \lim_{n\to\infty} \frac{n(n-1)\dots(n-x+1)}{x!} \left(\lambda/n\right)^x \left(1-\lambda/n\right)^{n-x}$$

$$\Rightarrow \lim_{n\to\infty} \frac{1(1-1/n)(1-2/n)\dots(1-x-1/n)}{x!} \times \lambda^x \left[(1-\lambda/n)^{-n/\lambda}\right]^{-\lambda} (1-\lambda/n)^{-x}$$

$$\lim_{f(x)\to 0} [(1+f(x)]^{1/f(x)} = e$$

$$\Rightarrow p(x) = 1/x! \times \lambda^x \times e^{-\lambda}$$

در نتیجه متغیر تصادفی x تمام خصوصیات تابع احتمال را دارد و تابع توزیع آن تقریبی از تابع توزیع متغیر تصادفی دو جمله ای است.

توزیع پواسون ادامه

$$E(x) = np = \lambda$$
$$\mathrm{var}(x) = npq = \lim_{n\to\infty} \lambda\left(1 - \lambda/n\right) = \lambda$$

$$E(x) = \sum_{x=0}^{\infty} x\frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda}\sum_{x=1}^{\infty}\frac{\lambda^x}{(x-1)!} = e^{-\lambda}\sum_{k=0}^{\infty}\frac{\lambda^{k+1}}{k!} = e^{-\lambda}\times\lambda\sum_{k=0}^{\infty}\frac{\lambda^k}{k!} = e^{-\lambda}\times\lambda\times e^{\lambda} = \lambda$$
$$Var(x) = ???$$

چگونگی ساخت عدد تصادفی با خصوصیت توزیع پواسون

مثال

اگر تقاضا در مهلت تحویل برای محصولی دارای توزیع پواسون با میانگین ۱۰ داشته باشد، با فاصله اطمینان ۹۵٪ در برابر کمبود نقطه سفارش مجدد را مشخص نمایید.

$$p(x) = \frac{e^{-\lambda}\lambda^x}{x!} = \frac{e^{-10}10^x}{x!}$$

$$\sum_{i=0}^{x}\frac{e^{-10}10^i}{i!} \ge 0.95$$
$$x = 1, x = 2, \ldots. x = 15$$

Prepared By Dr. Kazemipoor

توزیع های رایج متغیرهای تصادفی پیوسته

توزیع یکنواخت

- متغیر تصادفی یکنواخت X در بازه S=[a,b],b>a مقادیری را اختیار می کند که دارای احتمال یکسان می باشند. توزیع یکنواخت به صورت Unif(a,b) نشان داده می شود.

تابع چگالی به شکل زیر می باشد:

$$f(x)=\begin{cases}\dfrac{1}{b-a} & a\le x\le b\\ 0 & otherwise\end{cases}\qquad \Rightarrow F(x)=\begin{cases}0 & x<a\\ \dfrac{x-a}{b-a} & a\le x\le b\\ 1 & x>b\end{cases}$$

$$p(x_1<X<x_2)=\frac{x_2-x_1}{b-a}\qquad E(x)=\frac{a+b}{2}$$

$$\operatorname{var}(x)=\frac{(b-a)^2}{12}$$

چگونگی ساخت عدد تصادفی با خصوصیت توزیع یکنواخت پیوسته

توزیع مثلثی

متغیر تصادفی توزیع مثلثی X مقادیر موجود در بازه S=[a,b] را اختیار می کند. احتمال در زیربازه [a,c] به صورت خطی افزایش می یابد و در زیربازه [c,b] به صورت خطی کاهش می یابد. بنابراین تابع چگالی این متغیر دارای شکل مثلثی می باشد. توزیع مثلثی با نماد Tria(a,c,b) نشان می دهند و تابع چگالی آن به صورت زیر به دست می آید:

$$\Rightarrow f(x)=\begin{cases}\dfrac{2(x-a)}{(b-a)(c-a)} & a\le x\le c\\ \dfrac{2(b-x)}{(b-a)(b-c)} & c\le x\le b\\ 0 & x>b\end{cases}\qquad \Rightarrow\begin{cases}E(x)=\dfrac{a+b+c}{3}\\ Var(x)=\dfrac{a^2+b^2+c^2-ab-ac-bc}{18}\end{cases}$$

چگونگی ساخت عدد تصادفی با خصوصیت توزیع نمایی

توزیع نمایی

$$x\sim Exp(\lambda)$$

$$f(x)=\begin{cases}\lambda e^{-\lambda x} & x\ge 0\\ 0 & otherwise\end{cases}\qquad \Rightarrow F(x)=1-e^{-\lambda x}x\ge 0$$

$$E(x)=1/\lambda$$

$$\operatorname{var}(x)=1/\lambda^2$$

$$p(x>s+t\mid x>s)=p(x>t)$$ **خاصیت بی حافظگی توزیع نمایی**

$$\frac{p(x>s+t)}{p(x>s)}=\frac{e^{-\lambda(s+t)}}{e^{-\lambda s}}=e^{-\lambda t}=p(x>t)$$

چگونگی ساخت عدد تصادفی با خصوصیت توزیع نمایی

توزیع گاما

- تابع گاما

$$\Gamma(B)=\int_0^\infty x^{\beta-1}e^{-x}dx=(\beta-1)\Gamma(\beta-1)$$
$$\beta\in \mathrm{N}\Rightarrow\Gamma(\beta)=(\beta-1)!$$

- تابع توزیع گاما

$$x\sim Gamma(\beta,\lambda)$$
$$f(x)=\begin{cases}\dfrac{\lambda^\beta}{\Gamma(\beta)}x^{\beta-1}e^{-\lambda x} & x>0\\ 0 & otherwise\end{cases}\qquad \beta=1\Rightarrow f(x)=\frac{\lambda'}{\Gamma(1)}x^{1-1}e^{-\lambda x}=\lambda e^{-\lambda x}$$

توزیع گاما ادامه

If X_i has $Exp(\lambda)$ distributaion then $X=\sum_{i=1}^{\beta}X_i$ has $Gamma(\beta,\lambda)$ distributaion

$$\Rightarrow\begin{cases}E(x)=E\left(\sum_{i=1}^{\beta}X_i\right)=\sum_{i=1}^{\beta}E(X_i)=\beta*E(X_i)=\dfrac{\beta}{\lambda}\\ Var(x)=Var\left(\sum_{i=1}^{\beta}X_i\right)=\sum_{i=1}^{\beta}Var(X_i)=\beta*Var(X_i)=\dfrac{\beta}{\lambda^2}\end{cases}$$

چگونگی ساخت عدد تصادفی با خصوصیت توزیع گاما

توزیع کای دو

$$x \sim Gamma(\beta = \frac{v}{2}, \lambda = \frac{1}{2})$$

$$f(x) = \begin{cases} \dfrac{(\frac{1}{2})^{\frac{v}{2}}}{\Gamma\left(\frac{v}{2}\right)} x^{\frac{v}{2}-1} e^{-\frac{x}{2}} & x > 0 \\ 0 & otherwise \end{cases}$$

$$\Rightarrow \begin{cases} x \sim \text{Chi-square}(x) \\ f(x) = \begin{cases} \dfrac{(\frac{1}{2})^{\frac{v}{2}}}{\Gamma\left(\frac{v}{2}\right)} x^{\frac{v}{2}-1} e^{-\frac{x}{2}} & x > 0 \\ 0 & otherwise \end{cases} \end{cases}$$

چگونگی ساخت عدد تصادفی با خصوصیت توزیع کای دو

توزیع ارلنگ

- همان تابع توزیع گاما در حالتی است که $\beta \in N$ است. در این حالت داریم

$$x \sim Erlang(\beta, \lambda)$$

$$f(x) = \frac{\lambda^{\beta}}{(\beta-1)!} x^{\beta-1} e^{-\lambda x} \Rightarrow F(x) = 1 - e^{-\lambda x} \sum_{i=0}^{\beta-1} \frac{(\lambda x)^i}{i!}$$

چگونگی ساخت عدد تصادفی با خصوصیت توزیع ارلنگ

مثال

دو لامپ با عمر متوسط ۱۰۰۰ ساعت با توزیع نمایی به گونه‌ای بسته شده‌اند که در صورت خارج شدن یکی لامپ دیگر روشن می‌شود. احتمال اینکه بعد از ۲۱۶۰ ساعت لامپی روشن باشد، چقدر است.

- $X=X_1+X_2$
  - X_1 ~EXP (1/1000)
  - X_2~EXP (1/1000)

$$F(x)=1-e^{-\lambda x}\sum_{i=0}^{\beta-1}\frac{(\lambda x)^i}{i!}=1-e^{-\frac{1}{1000}x}\sum_{i=0}^{1}\frac{\left(\frac{1}{1000}x\right)^i}{i!}$$

$$F(x)=1-e^{-2.160}\sum_{i=0}^{1}\frac{(2.16)^i}{i!}=0.636=0.64$$

$$1-0.64=0.36$$

مثال

یک معاینه پزشک سه مرحله دارد که هر کدام دارای توزیع نمایی با میانگین ۲۰ دقیقه است. با این فرضیات احتمال مدت معاینه کمتر از ۵۰ دقیقه باشد را بیابید.

$X=X_1+X_{2+}X_3$

- X_1 ~EXP (1/20)
- X_2~EXP (1/20)
- X_3~EXP (1/20)

$$F(50)=1-e^{-\frac{1}{20}\times 50}\times\sum_{i=0}^{2}\frac{\left(\frac{1}{20}\times 50\right)^i}{i!}=1-e^{-2.5}\sum_{i=0}^{2}\frac{(2.5)^i}{i!}=1-0.543=0.457$$

$$E(x)=\frac{k}{\lambda}=\frac{3}{\frac{1}{20}}=60 \qquad \hat{M}=\frac{k-1}{\lambda}=\frac{3-1}{\frac{1}{20}}=\frac{2}{\frac{1}{20}}=40$$

توزیع نرمال

$$X \sim N(\mu,\sigma^2)$$

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad -\infty < x < \infty$$

$$1)\lim_{x\to+\infty} f(x)=\lim_{x\to-\infty} f(x)=0$$

$$2)\ f(x+\mu)=f(x-\mu)$$

$$3)\ \text{Arg}[Max(f(x))]=\mu$$

$$4)\ F(x)=p(X<x)=\int_{-\infty}^{x}\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}dt=\int_{-\infty}^{\frac{x-\mu}{\sigma}}\frac{1}{\sqrt{2\pi}}e^{\frac{-1}{2}z^2}dz=\Phi\left(\frac{x-\mu}{\sigma}\right)$$

$$5)\ f(z)=\frac{1}{\sqrt{2\pi}}e^{\frac{-1}{2}z^2};-\infty<z<+\infty\to z\sim N(0,1)$$

امید ریاضی و واریانس توزیع نرمال

$$X \sim N(\mu,\sigma^2)$$

$$E(x)=\int_{-\infty}^{+\infty}xf(x)d_x=\int_{-\infty}^{+\infty}x\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}d_x=\mu$$

$$Var(x)=E(x^2)-[E(x)]^2=(\mu^2+\sigma^2)-(\mu)^2=\sigma^2$$

چگونگی ساخت عدد تصادفی با خصوصیت توزیع نرمال

مثال

If $X \sim N(50,9)$ Calculate $p(x \le 56)$

$$p(x \le 56) = p\left(\frac{X-50}{3} < \frac{56-50}{3}\right) = p(z<2) = \Phi(2) = 0.9772$$

If $X \sim N(12,4)$ Calculate $p(x \le 10)$ and $p(10 \le x \le 12)$

$$F(10) = \Phi\left(\frac{10-12}{2}\right) = \Phi(-1) = 0.1587$$

$$p(10 \le X \le 12) = F(12) - F(10) = \Phi\left(\frac{12-12}{2}\right) - \Phi\left(\frac{10-12}{2}\right) = \Phi(0) - \Phi\left(-2/2\right) = \Phi(0) - (1-\Phi(+1))$$

$$= 0.5 - (1-\Phi(+1)) = 0.5 - (1-0.084134) = 0.3413$$

مثال

اگر تقاضا برای محصولی دارای توزیع نرمال با میانگین ۲۵ و واریانس ۹ باشد، نقطه سفارش مجدد را به گونه‌ای بیابید که کمبود فقط در ۵٪ مواقع رخ دهد.

If $X \sim N(50,9)$ Calculate $p(x \le 56)$

$$p(x > x_0) = 0.05$$

$$1 - \Phi\left(\frac{x_0 - 25}{3}\right) = 0.05$$

$$\Phi\left(\frac{x_0 - 25}{3}\right) = 0.95 \Rightarrow \frac{x_0 - 25}{3} = 1.65 \Rightarrow 29.935 \cong 30$$

اگر به هنگام رسیدن تقاضا به ۳۰ واحد سفارش خرید صادر شود فقط در ۵٪ مواقع کمبود داریم.

توزیع لوگنرمال

$X \sim Logn(\mu, \sigma^2)$

$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} e^{\frac{(\ln x - \mu)}{2\sigma^2}} \quad , \quad x \geq 0$$

$$E(x) = e^{\mu + \frac{\sigma^2}{2}}$$

$$Var(x) = e^{2\mu + \sigma^2}(e^{\sigma^2} - 1)$$

چگونگی ساخت عدد تصادفی با خصوصیت توزیع لوگنرمال

متغیر تصادفی لوگنرمال همواره مثبت می باشد و اغلب برای مدلسازی فرآیندهای تصادفی مالی به کا می رود.

Note

If Y hase Lognormal Distribution and $Y = e^X$ then X has Normal Distribution

توزیع بتا

$$\text{BETA Function}: B(\alpha, \beta) = \int_0^\infty x^{\alpha-1}(1-x)^{\beta-1} d_x = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

$$\text{BETA Distribution}: f(x) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{B(\alpha, \beta)} \quad , \quad 0 \leq X \leq 1$$

$$E(x) = \frac{\alpha}{\alpha+\beta}$$

$$Var(x) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

یک متغیر تصادفی بتا، اغلب در آمار برای مدلسازی یک احتمال نامعلوم که متغیر تصادفی نامیده می شود به کار می رود.

چگونگی ساخت عدد تصادفی با خصوصیت توزیع بتا

توزیع وایبول

$$X \sim Weibull(\alpha, \beta)$$

$$f(x) = \frac{\alpha}{\beta^{\alpha}} x^{\alpha-1} e^{-(\frac{\alpha}{\beta})^{\alpha}} \quad , \ 0 \le X$$

$$E(x) = \alpha \Gamma(\frac{1}{\beta} + 1)$$

متغیرهای تصادفی وایبول اغلب در مدلسازی فرآیند فرسودگی اجزا در تحلیل قابلیت اطمینان استفاده می شوند.

$$Var(x) = \alpha^2 \left[\Gamma(\frac{2}{\beta} + 1) - \Gamma^2(\frac{1}{\beta} + 1) \right]$$

چگونگی ساخت عدد تصادفی با خصوصیت توزیع وایبول

توزیع تی استیودنت

$$X \sim T - Student(n)$$

$$f_X(x) = \frac{\Gamma((n+1)/2)}{\sqrt{\pi n}\ \Gamma(n/2)} \left(1 + \frac{x^2}{n}\right)^{-(n+1)/2}, \quad -\infty \le x \le \infty$$

$$E(x) = 0 \qquad , \qquad Var(x) = \frac{n}{n-2}$$

Where Z is a standard normal random variable, Y is a chi-square random variable with n degrees of freedom, and Z and Y are independent. X has T-Student distribution

$$X = \frac{Z}{\sqrt{Y/n}}$$

چگونگی ساخت عدد تصادفی با خصوصیت توزیع تی استیودنت

توزیع فیشر

$$X \sim F(n_1, n_2)$$

$$f_X(x) = \frac{\Gamma((n_1+n_2)/2)}{\Gamma(n_1/2)\Gamma(n_2/2)} \left[\frac{n_1}{n_2}\right]^{n_1/2} \frac{x^{(n_1/2)-1}}{\left[1+\frac{n_1}{n_2}x\right]^{(n_1+n_2)/2}}, \quad 0 \le x \le \infty$$

$$\mathrm{E}[X] = \frac{n_2}{n_2-2} \quad (\text{for } n_2 > 2)$$

$$\mathrm{V}[X] = \frac{2n_2^2(n_1+n_2-2)}{n_1(n_2-4)(n_2-2)} \quad (\text{for } n_2 > 4).$$

where V and W are independent chi-square random variables with the corresponding degrees of freedom n_1,n_2 X has F distribution with n_1,n_2 degrees of freedom.

$$X = \frac{V/n_1}{W/n_2}$$

چگونگی ساخت عدد تصادفی با خصوصیت توزیع فیشر

نکته‌ای در باره توزیع‌های رایج آماری

می‌توان بسیاری از اتفاقاتی که در پیرامون ما رخ می‌دهد را به یکی از توزیع‌های گفته شده با فاصله اطمینان خاصی نسبت داد. پس اگر بتوانیم در مورد یک متغیر تصادفی، یکی از این توزیع‌ها برازش کنیم، تحلیل‌ها در شبیه‌سازی سمت و سوی مشخص تری می‌گیرد. در ادامه این فصل تعیین توزیع مناسب و تخمین پارامترها برای یک سری ازداده‌های مشخص انجام خواهد گرفت.

Poisson Process

A Poisson process is a special type of counting process that is a fundamental base case for defining many other types of counting processes.

Definition: The counting process $\{N(t), t \geq 0\}$ is said to be a Poisson process with rate λ, $\lambda > 0$, if

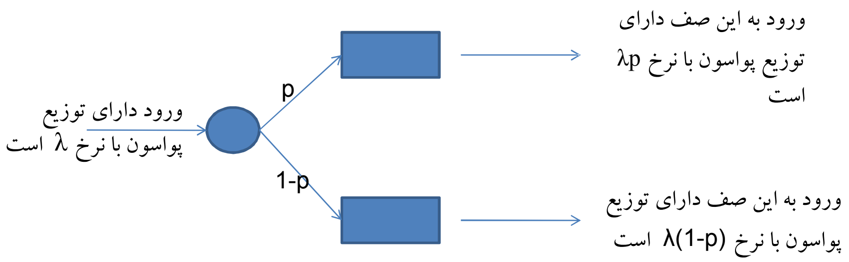
- $N(0) = 0$
- the process has independent increments
- the number of events in any interval of length t is Poisson distributed with mean λt.

Poisson Process

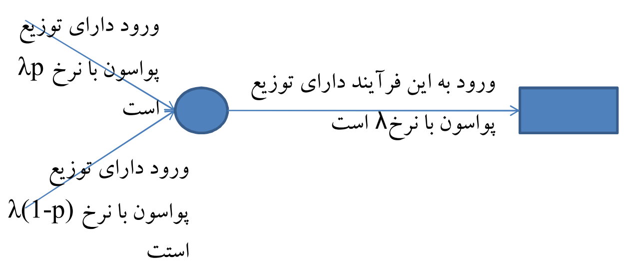
The single parameter λ controls the rate at which events occur over time. Since λ is a constant, a Poisson process is often referred to as a homogeneous Poisson process. The third condition is equivalent to

$$p(N(t+s)-N(s)=n)=\frac{e^{-\lambda t}(\lambda t)^n}{n!} \quad ; \quad n=0,1,\ldots$$

نکاتی در توزیع پواسون



نکاتی در توزیع پواسون ادامه



مثال

آهنگ ورود به یک فرایند دارای توزیع پواسان با میانگین ۲۰ دقیقه است. احتمال اینکه در یک دوره ۲ ساعته هیچ سفارش وارد نشود چیست؟

$$\mu = 20 \;\min \Rightarrow \lambda = {}^{60}\!/_{20} = 3$$

$$\lambda t = 3\times 2 = 6 \Rightarrow p(0) = \frac{e^{-6}6^{0}}{0!} = 0.003$$

رابطه میان فرآیند پواسون و توزیع نمایی

اگر تعداد پیروزی ها در فاصله زمانی ۰ تا t دارای توزیع پواسون باشد، می توان نشان داد که زمان رسیدن به اولین پیروزی دارای توزیع نمایی است.

اثبات:

$$X \sim Poisson(\lambda t)$$

$$\Rightarrow p(N(t)=n) = \frac{e^{-\lambda t}(\lambda t)^n}{n!} \quad ; \quad n = 0,1,\dots$$

$$p(T \le t) = 1 - p(T > t) = 1 - p(\text{صفر پیروزی در فاصله زمانی صفر تا } t) =$$

$$= 1 - \frac{e^{-\lambda t}(\lambda t)^0}{0!} = 1 - e^{-\lambda t} \Rightarrow f(t) = \frac{\partial}{\partial t}(1 - e^{-\lambda t}) = \lambda e^{-\lambda t}$$

$$\Rightarrow t \sim Exp(\lambda)$$

ساخت اعداد تصادفی

عدد تصادفی

یکی از موارد کلیدی در شبیه سازی سیستم های گسسته پیشامد تولید اعداد شبه تصادفی است. بدین منظوردو مرحله کلی وجود دارد:

– ساخت مجموعه ای از اعداد

– تست این که اعداد تولید شده تصادفی اند، یعنی

- دارای توزیع یکنواخت بین صفر تا یک باشند (یعنی احتمال قرار گرفتن در هر فاصله‌ای برابر با طول آن فاصله باشد).
- اعداد تولید شده مستقل از هم باشند (یعنی هیچ‌گونه ارتباطی بین مقدار فعلی متغیر تصادفی و مقدار پیشین آن وود نداشته باشد).

خواص ضمنی اعداد تصادفی

- تابع چگالی
- امید ریاضی و واریانس
- اگر فاصله (۱و۰) به n رده یا زیر فاصله مساوی تقسیم شود، انتظار می‌رود که از N مشاهده N/n در هر رده قرار گیرد.
- احتمال مشاهده یک عدد در یک رده فاصله خاص مستقل از سایر مشاهده‌هاست. عنی همبستگی بین اعداد وجود نداشته باشد.

خواص اعداد تصادفی

- روش یا الگوریتم تولید اعداد تصادفی می‌بایست سریع باشد.
- الگوریتم نباید نیاز به مقدار زیادی حافظه کامپیوتر داشته باشد و می‌بایست قابل برنامه‌نویسی کامپیوتری باشد.
- طول دنباله اعداد تولید شده باید به اندازه کافی بلند باشد (به دلیل اینکه در نهایت از یک الگوریتم برای تولید اعداد تصادفی استفاده می‌شود. ایجاد سیکل اجتناب ناپذیر خواهد بود ولی طول سیکل بلند (مثلا چند میلیون و یا چند میلیارد) اهداف شبیه‌سازی را تأمین خواهد کرد).

Prepared By Dr. Kazemipoor

روش های تولید اعداد تصادفی

- روش میان ضربی
- روش مضرب ثابت
- روش همنهشتی جمعی
- مولدهای همنهشتی خطی liner congruential method
- مولدهای همنهشتی خطی ترکیبی Combined liner congruential method

liner congruential method

- X_0: مقدار اولیه هسته
- a: ضریب ثابت مولد
- c: مقدار ثابت مولد
- m: مقدار پیمانه

$$X_1 \overset{m}{\equiv} \left(aX_{i-1} + c\right)$$

نحوه انتخاب مقادیر پارامترها تأثیر فراوانی در خواص آماری از قبیل میانگین، واریانس و طول سیکل دارد. وقتی c مخالف صفر باشد مولد رامولد همنهشتی آمیخته می‌نامند و در صورتی که c برابر صفر باشد مولد همنهشت ضربی نامیده می‌شود.

مثال

با استفاده از مقادیر x$_0$=27, a=17, c=43, m=100 و بر اساس روش همنهشتی خطی دنباله‌ای از اعداد تصادفی تولید کنید.

$$X_0 = 27$$
$$X_1 \overset{100}{\equiv} [17(27)+43] \overset{100}{\equiv} 502 \overset{100}{\equiv} 02$$
$$R_1 = 0.02$$
$$X_2 \overset{100}{\equiv} [17(2)+43] \overset{100}{\equiv} 77 \overset{100}{\equiv} 77$$
$$R_1 \equiv 0.77$$
$$X_3 \overset{100}{\equiv} [17(77)+43] \overset{100}{\equiv} 1352 \overset{100}{\equiv} 52$$
$$R_3 = 0.52$$

Example

For a concrete illustration, let $m = 2^{31} - 1 = 2147483647$, $c = 0$, and $a = 16807$. These parameters were originally proposed in [83]. Take $x_0 = 12345$. Then

$$x_1 = 16,807 \times 12,345 \bmod m = 207,482,415$$
$$u_1 = \frac{207,482,415}{m} = 0.0966165285$$
$$x_2 = 16,807 \times 207,482,415 \bmod m = 1,790,989,824$$
$$u_2 = \frac{1,790,989,824}{m} = 0.8339946274$$
$$x_3 = 16,807 \times 1,790,989,824 \bmod m = 2,035,175,616$$
$$u_3 = \frac{2,035,175,616}{m} = 0.9477024977$$

Combined liner congruential method

با ترکیب دو یا چند مولد همنهشتی ضربی ممکن است طول دنباله اعداد تولید شده بیشتر شده و در نتیجه احتمال تولید اعداد بهتری وجود داشته باشد.

تست یکنواختی و تست های استقلال

۱- آزمون های فراوانی برای تست یکنواختی توزیع اعداد تولید شده

۱-۱- آزمون مربع کای

۱-۲- آزمون کالموگروف-اسمیرنف

۲- آزمون های استقلال اعداد تولید شده

۲-۱- آزمون روند

۲-۱-۱- روندهای صعودی و نزولی

۲-۱-۲- روندهای کوچکتر و بزرگتر از میانگین

۲-۱-۳- آزمون روند براساس طول روند (طول روندهای بزرگتر و کوچکتر از میانگین، طول روندهای صعودی و نزولی

۲-۲- آزمون همبستگی

۲-۳- آزمون افراز

۲-۴- آزمون شکاف

تست یکنواختی توزیع اعداد تولید شده

الگوریتم های تست یکنواختی اعداد تصادفی بر پایه تئوری‌های آماری، یا آزمونهای فرض می‌باشند. به عنوان مثال در تست توزیع یکنواخت ما دو فرض داریم که یکی بیان می کند که اعداد تصادفی توزیع یکنواخت دارند که ما آن را فرض صفر می‌نامیم و دیگری بیان می کند که اعداد تصادفی توزیع یکنواخت ندارند و ما ان را H_1 می‌نامیم که در آمار به عنوان فرض جایگزین شناخته می‌شود. در این تست آماری علاقه‌مند به بررسی نتیجه فرض صفر رد کردن آن و یا عدم رد آن هستیم.

$H_0 : R_i \sim U[0,1]$ اعداد توزیع یکنواخت دارد

$H_1 : R_i \sim \cancel{U[0,1]}$ اعداد دارای توزیع یکنواخت نیستند

(صحت فرض صفر|رد فرض صفر)p=احتمال خطای نوع اول

(فرض صفر غلط|پذیرش فرض صفر) p=احتمال خطای نوع دوم

مثال ۱- تست یکنواختی KS

فرض کنید توسط یک رویه ساخت اعداد تصادفی، اعداد زیر ساخته شده اند:

۰.۸۱ و ۰.۱۴ و ۰.۰۵ و ۰.۹۳ و ۰.۴۴

آیا این اعداد به طور یکنواخت توزیع شده اند؟

$R_{(i)}$	0.05	0.14	0.44	0.81	0.93
i/N	0.20	0.40	0.60	0.80	1.00
$i/N - R_{(i)}$	0.15	0.26	0.16	—	0.07
$R_{(i)} - (i-1)/N$	0.05	—	0.04	0.21	0.13

$D^+ = 0.26$

$D^- = 0.26$

$$\Rightarrow \left\{ \begin{array}{l} D = Max\{0.21, 0.26\} = 0.26 \\ D_{0.05,5} = 0.56 \end{array} \right\} \Rightarrow D_{0.05,5} > D \Rightarrow$$ دلیلی بر رد فرض صفر وجود ندارد

مثال ۲- آزمون کای دو

اعداد زیر با یک رویه خاص تولید شده اند، آیا این اعداد در سطح معنی دار ۹۵٪ دارای تابع توزیع یکنواخت می باشند؟

0.34	0.90	0.25	0.89	0.87	0.44	0.12	0.21	0.46	0.67
0.83	0.76	0.79	0.64	0.70	0.81	0.94	0.74	0.22	0.74
0.96	0.99	0.77	0.67	0.56	0.41	0.52	0.73	0.99	0.02
0.47	0.30	0.17	0.82	0.56	0.05	0.45	0.31	0.78	0.05
0.79	0.71	0.23	0.19	0.82	0.93	0.65	0.37	0.39	0.42
0.99	0.17	0.99	0.46	0.05	0.66	0.10	0.42	0.18	0.49
0.37	0.51	0.54	0.01	0.81	0.28	0.69	0.34	0.75	0.49
0.72	0.43	0.56	0.97	0.30	0.94	0.96	0.58	0.73	0.05
0.06	0.39	0.84	0.24	0.40	0.64	0.40	0.19	0.79	0.62
0.18	0.26	0.97	0.88	0.64	0.47	0.60	0.11	0.29	0.78

Interval	O_i	E_i	$O_i - E_i$	$(O_i - E_i)^2$	$\frac{(O_i - E_i)^2}{E_i}$
1	8	10	−2	4	0.4
2	8	10	−2	4	0.4
3	10	10	0	0	0.0
4	9	10	−1	1	0.1
5	12	10	2	4	0.4
6	8	10	−2	4	0.4
7	10	10	0	0	0.0
8	14	10	4	16	1.6
9	10	10	0	0	0.0
10	11	10	1	1	0.1
	100	100	0		3.4

$$\chi^2_{0.05,9} = 16.9 \Rightarrow$$

دلیلی بر رد فرض

صفر وجود ندارد

مقایسه آزمون مربع کای و کالموگروف- اسمیرنف

- آزمون k-s تک تک مشاهدات را در نظر می‌گیرد، ولی کای دو مشاهدات را رده بندی کرده و بدین ترتیب تعدادی از داده‌ها حذف شده و دقت کم می‌شود.
- در مواردی که تعداد مشاهدات کم است، k-s به دلیل دقیق بودن قابل اعمال است، ولی کای دو بیشتر برای نمونه‌های بزرگ کاربرد دارد.
- در کای دو امکان برآورد پارامترها از طریق مشاهده نیز وجود دارد (با تغییر آزمون) ولی k-s این انعطاف‌پذیری را ندارد.
- کای دو هم در مورد داده‌های پیوسته و هم گسسته قابل به کارگیری است، ولی k-s تنها برای مواردی که تابع توزیع تجمعی جهشی نیست. قابل به کارگیری است.

آزمون همبستگی

اگر داده های تولید شده توسط یک مولد خاص، تصادفی باشند، نبایستی هیچ نظمی به هیچ طریقی میان آن ها وجود داشته باشد. در حقیقت برای بررسی عدم وجود نظم میان داده ها، آزمون فرض زیر انجام می شود.

$$H_0 : \rho_{im} = 0$$
$$H_1 : \rho_{im} \neq 0$$

در این آزمون می توان نشان داد که آماره آزمون برابر است با: $Z = \dfrac{\hat{\rho}_{im}}{\hat{\sigma}_{\hat{\rho}_{im}}}$

در این رابطه پارامترهای به صورت زیر به دست می آید:

$$\hat{\rho}_{im} = \frac{1}{M+1}\left[\sum_{k=0}^{M} R_{i+km} R_{i+(k+1)m}\right] - 0.25 \qquad \hat{\sigma}_{\hat{\rho}_{im}} = \frac{\sqrt{13M+7}}{12(M+1)}$$

مثال

توسط یک مولد خاص، اعداد زیر تولید شده است. آیا بر اساس تست همبستگی، می توان اعداد سوم، هشتم، سیزدهم، ... را در سطح ۹۵٪ تصادفی دانست؟

۰.۱۲	۰.۰۱	۰.۲۳	۰.۲۸	۰.۸۹	۰.۳۱	۰.۶۴	۰.۲۸	۰.۸۳	۰.۹۳
۰.۹۹	۰.۱۵	۰.۳۳	۰.۳۵	۰.۹۱	۰.۴۱	۰.۶۰	۰.۲۷	۰.۷۵	۰.۸۸
۰.۶۸	۰.۴۹	۰.۰۵	۰.۴۳	۰.۹۵	۰.۵۸	۰.۱۹	۰.۳۶	۰.۶۹	۰.۸۷

حل

$i = 3 \qquad N = 30 \qquad M = 4$

$\alpha = 0.05 \Rightarrow Z_{\frac{0.05}{2}} = 1.96$

$$\hat{\rho}_{im} = \frac{1}{4+1}(0.23*0.28 + 0.28*0.33 + 0.33*0.27$$
$$+ 0.27*0.05 + 0.05*0.36) - 0.25 = -0.1945$$

$$\hat{\sigma}_{\hat{\rho}_{im}} = \frac{\sqrt{13*4+7}}{12(4+1)} = 0.1280$$

$$\Rightarrow Z = \frac{-0.1945}{0.1280} = -1.516 \Rightarrow -1.96 \le -1.516 \le 1.96$$

دلیلی بر رد فرض صفر وجودندارد

تحلیل داده‌های ورودی

Data Collection

Data are of course central to the development and use of simulation models. Much effort may go into the design of the conceptual model and into the coding of the computer model. However, if the data that are used to design and populate the model are inaccurate then the results from the model will also be inaccurate. The concentration on quantitative data, however, ignores the importance of qualitative data as well.

طراحی مدل

به منظور طراحی مدل معتبر از داده‌های ورودی مسله چهار قدم ضروری است:

1. گردآوری داده‌های خام
2. تعیین توزیع‌های آماری برای داده‌های خام
3. برآوردهایی پارامترهای مشخص کننده توزیع
4. آزمون توزیع‌ها

– در پایان این چهار مرحله در صورت رد فرض صفر به قدم دوم بر می‌گردیم.

۱- گردآوری داده‌های خام

- برنامه‌ریزی: طراحی برگه‌هایی برای جمع‌آوری اطلاعات با توجه به شرایط قبلی مسأله.
  - البته در مراحل اولیه اشکالی ندارد که این برگه‌ها چند بار تصحیح شود. این نکته را در جمع آوری اطلاعات به یاد داشته باشید که همواره در پی شناسایی اوضاع و احوال غیر معمول پیرامون مسأله باشد.
- تجزیه و تحلیل همزمان با گردآوری داده‌ها
  - کافی بودن داده‌های گردآوری شده را از لحاظ مشخص کردن توزیع‌های آماری مورد نیاز به عنوان ورودی شبیه‌سازی تعیین کنید و داده‌های غیر مفیدی جمع‌آوری نکنید.
- ادغام مجموعه های همگن در داده ها
  - همگنی داده‌ها را در دوره‌های متوالی در چند روز مورد بررسی قرار دهید، مثلا داده‌های ۲-۳ روز شنبه با ۲-۳ پنج شنبه ادغام کنید. برای داغامی توانید آزمون‌های مقدماتی بررسی میانگین را انجام دهید.
- بررسی روابطه میان دو یا چند متغیر
- بررسی خود همبستگی داده‌های ظاهرا مستقل

۲- تعیین توزیع‌های آماری برای داده‌ها

۱- نمودار فراوانی

- تقسیم‌بندی داده‌های جمع‌آوری شده
- قسمت‌بندی محور افقی با استفاده از تقسیم‌بندی متناسب
- مشخص کردن فراوانی برای هر طبقه
- تقسیم‌بندی محور عمودی به گونه‌ای که همه فراوانی‌ها را در برگیرد
- رسم فراوانی‌ها

تعداد فاصله رده‌ای k	n
امکان پذیر نیست	۲۰
۵ تا ۱۰	۵۰
۱۰ تا ۲۰	۱۰۰
n/5	>100

۲- تعیین توزیع احتمال فرضی

- استفاده از توزیع های تجربی
- استفاده از توزیع های آماری معروف (میانگین، واریانس و ضریب تعیین توزیع می توانند کمک حال ما باشند)

We use Quantile- Quantile Plote for this section

۳- برآورد پارامترهای توزیع

برای برآورد پارامترهای یک توزیع بایستی از روشهای آماری برآورد پارامترها استفاده نمود. از جمله این روش ها می توان به موارد زیر اشاره نمود:

– روش گشتاورها

$$\left.\begin{aligned} m_k &= \frac{\sum x_i^k}{n} \\ \mu_k &= \sum x_i^k p(X = x_i) \end{aligned}\right\} \Rightarrow m_k = \mu_k$$

– روش تخمین ماکزیمم درستنمایی

– روش برآورد بیزی

براورد پارامترها در برخی توزیع ها

توزیع احتمال	پارامتر(ها)	برآوردکننده(های) پیشنهادی
پواسون	α	$\hat{\alpha} = \overline{X}$
نمایی منفی	λ	$\hat{\lambda} = 1/\overline{X}$
گاما	β, θ	$M = \ln \overline{X} - \frac{1}{n}\sum_{i=1}^{n} \ln X_i \Rightarrow \hat{\beta}$ $\hat{\theta} = 1/X$
یکنواخت در فاصله (0,b)	b	$\hat{b} = \frac{n+1}{n}[\max X]$
نرمال	μ, σ^2	$\hat{\mu} = \hat{X}$ $\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - X)^2$
وایبول	α, β	$\hat{\beta}_j = X/S$ $\hat{\beta}_j = \hat{\beta}_{j-1} - \left(f\left(\hat{\beta}_{j-1}\right)/f'\left(\hat{\beta}_{j-1}\right)\right)$ $\hat{\alpha} = \left(\frac{1}{n}\sum_{i=1}^{n} X^{\hat{\beta}}{}_i\right)^{1/\hat{\beta}}$

٤- آزمون های برازندگی برای آزمون توزیع‌ها

- آزمون مربع کای
- آزمون مربع کای با احتمال های برابر
- آزمون برازندگی کولموگروف- اسمیرنف KS

مثال ۱- آزمون مربع کای

فرض کنید با استفاده از نمونه گیری تعداد وسایل نقلیه وارد شده به یک تقاطع به صورت زیر به دست آمده است.

تعداد وسایلی که طی دوره ۵ دقیقه‌ای به تقاطع می‌رسند (X)	۰	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱
فراوانی	۱۲	۱۰	۱۹	۱۷	۱۰	۸	۷	۵	۵	۳	۳	۱

فرض کنید با مقایسه این جدول با جدول توزیع احتمال پواسون چنین به نظر رسیده است که این اعداد دارای توزیع احتمال فرضی پواسون باشد، آزمون مربوطه را برای تست این فرضیه انجام دهید.

مثال ۱ ادامه

اولین مرحله برای تست این فرضیه بدست آوردن پارامتر توزیع فرض شده برای اعداد داده شده است. با توجه به جدول برآوردهای ارائه شده برای توزیع های آماری داریم:

$$\hat{\lambda} = \overline{X} = \frac{\sum x_i}{n} = 3.64$$

H_0 : X has poisson distribution

H_1 : X Don't has poisson distribution

If x has poisson distribution then $p(x) = \frac{e^{-\lambda}\lambda^x}{x!} = \frac{e^{-3.84}3.84^x}{x!}$

Prepared By Dr. Kazemipoor

مثال ۱ ادامه

We can say that:

$$p(x=0) = \frac{e^{-3.84}3.84^0}{0!} = 0.026 \Rightarrow \text{Expected frequency}: 100*0.026 = 2.6$$

$$p(x=1) = \frac{e^{-3.84}3.84^1}{1!} = 0.096 \Rightarrow \text{Expected frequency}: 100*0.096 = 9.6$$

$$p(x=2) = \frac{e^{-3.84}3.84^2}{2!} = 0.174 \Rightarrow \text{Expected frequency}: 100*0.174 = 17.4$$

....

$$p(x=10) = \frac{e^{-3.84}3.84^{10}}{10!} = 0.003 \Rightarrow \text{Expected frequency}: 100*0.003 = 0.3$$

$$p(x=11) = \frac{e^{-3.84}3.84^{11}}{11!} = 0.001 \Rightarrow \text{Expected frequency}: 100*0.001 = 0.1$$

Prepared By Dr. Kazemipoor

مثال ۱ ادامه

X_i	فراوانی مشاهده شده O_i	فراوانی انتظاری	$\frac{(O_i\text{-}E_i)^2}{E_i}$
0	12	2.6	…
1	10	9.6	7.87
2	19	17.4	0.15
3	17	21.1	0.80
4	10	19.2	4.41
5	8	14	2.57
6	7	8.5	0.26
7	5	4.4	…
8	5	2	…
9	3	0.8	…
10	3	0.3	…
11,…	1	0.1	11.62

مثال ۱ ادامه

$$\left.\begin{aligned}\chi^2_{Sample} &= \sum_{i=1}^{7}\frac{(O_i\text{-}E_i)^2}{E_i} = 27.68\\ \chi^2_{0.5,7-1-1} &= \chi^2_{0.05,5} = 11.1\end{aligned}\right\}\chi^2_{Sample} \geq \chi^2_{0.05,5} \Rightarrow \text{We reject } H_0$$

آزمون کای با احتمال برابر

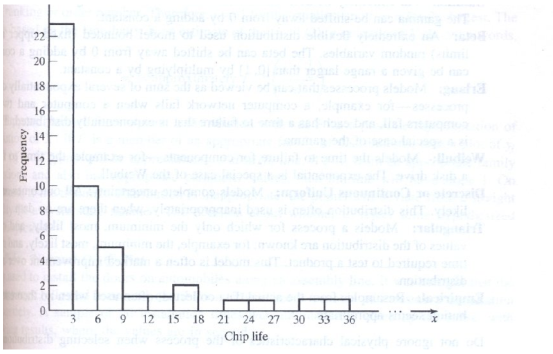
سوال اساسی در مواقعی که با توزیع های آماری پیوسته سر و کار داریم، اینست طول هر دسته در رده بندی داده هاچقدر باشد. یک راهکار برای این امر بدین طریق توصیه می شود که بازه هایی با احتمال های یکسان ایجاد کنیم. این امر در ادامه با مثال توضیح داده می شود

مثال ۲- آزمون کای با احتمال برابر برای توزیع نمایی

یک نمونه ۵۰ تایی از چیپ های نیمه رسانا مورد آزمایش تعیین عمر قرار گرفته اند. نتایج عمر بر حسب روز به شرح جدول زیر شده است.

۵.۸۴۵	۱.۹۶۱	۰.۰۶۲	۳.۰۸۱	۷۹.۹۱۹
۰.۱۲۳	۰.۰۱۳	۰.۰۲۱	۶.۵۰۵	۳.۰۲۷
۵.۰۰۹	۳۴.۷۶۰	۱.۱۹۲	۵۹.۸۹۹	۶.۷۶۹
۰.۴۳۳	۲۴.۴۲۰	۴۳.۵۶۵	۰.۱۴۱	۱۸.۳۸۷
۹.۰۰۳	۰.۰۹۱	۱۷.۹۶۷	۲.۶۶۳	۱۴۴.۶۹۵
۷.۵۷۹	۲.۱۵۷	۳.۳۷۱	۰.۸۷۸	۰.۹۴۱
۲۳.۹۶۰	۷.۰۷۸	۳.۱۴۸	۵.۳۸۰	۰.۶۲۴
۰.۵۴۳	۰.۰۰۲	۰.۳۰۰	۱.۹۲۸	۰.۵۹۰
۰.۲۱۹	۱.۱۴۷	۱.۰۰۵	۳۱.۷۶۴	۷.۰۰۴
۴.۵۶۲	۲.۳۳۶	۱.۰۰۸	۱۴.۳۸۲	۳.۲۱۷

مثال ۲- نمودار هیستوگرام



Prepared By Dr. Kazemipoor

مثال ۲- ادامه

$$F(a_i) = 1 - e^{-\lambda a_i} = iP \quad ; \quad i = 0,1,..,K$$

$$\Rightarrow e^{-\lambda a_i} = 1 - iP \Rightarrow a_i = -\frac{1}{\lambda}\ln(1 - ip) \quad ; \quad i = 0,1,..,K$$

We know that: $\hat{\lambda} = 0.084$ (Why?)

We recommend that: $np_i \geq 5$

If we have identical probability for each category then:

$$p_i = \frac{1}{k} \Rightarrow \frac{n}{k} \geq 5 \Rightarrow k \leq \frac{n}{5}$$

For this problem we suppose that $k = 8 \Rightarrow p = 0.125$

Prepared By Dr. Kazemipoor

مثال ۲– ادامه

$$a_0 = 0$$

$$a_1 = -\frac{1}{0.084}\ln(1-1*0.125) = 0.159$$

$$a_2 = -\frac{1}{0.084}\ln(1-2*0.125) = 3.425$$

....

$$a_7 = -\frac{1}{0.084}\ln(1-7*0.125) = 24.755$$

$$a_8 = \infty$$

مثال ۲– نتیجه

Class Interval	Observed Frequency, O_i	Expected Frequency, E_i	$\frac{(O_i - E_i)^2}{E_i}$
[0, 1.590)	19	6.25	26.01
[1.590, 3.425)	10	6.25	2.25
[3.425, 5.595)	3	6.25	0.81
[5.595, 8.252)	6	6.25	0.01
[8.252, 11.677)	1	6.25	4.41
[11.677, 16.503)	1	6.25	4.41
[16.503, 24.755)	4	6.25	0.81
[24.755, ∞)	6	6.25	0.01
	50	50	39.6

مثال ۳- تعیین فاصله رده ای در توزیع نرمال

$$F(x) = \Phi(\frac{x-\mu}{\sigma}) \quad ; \quad -\infty \le x \le \infty$$

$$F(a_1) = \Phi(\frac{a_1-\mu}{\sigma}) = 1*\mathrm{p}, \quad \text{If} \quad \mathrm{p} = 0.125 \Rightarrow \Phi(\frac{a_1-\mu}{\sigma}) = 0.125$$

$$\Rightarrow \Phi(\frac{a_1-\mu}{\sigma}) = \Phi(-1.152)$$

$$\Rightarrow a_1 = \mu - 1.152\sigma$$

آزمون KS

آزمون کالموگروف- اسمیرنف در مواردی که اندازه نمونه کوچک تر باشد، و هیچ یک از پارامترهای توزیع احتمال منتخب بر اساس داده های تجربی برآورد نشده باشد، این آزمون بسیار سودمند است.

تمرین

- امید و واریانس تابع توزیع هندسی را بدست آورید.
- امید و واریانس تابع توزیع گاما را به طور مستقیم بدست آورید.
- امید و واریانس تابع توزیع بتا را به طور مستقیم بدست آورید.
- امید و واریانس تابع توزیع وایبول را به طور مستقیم بدست آورید.
- با استفاده از یک نرم افزار آماری شکل تابع توزیع گاما، کای دو، وایبول و بتا را به ازای پارامترهای مختلف رسم نمایید.
- روش ساخت اعداد تصادفی دارای توزیع برنولی، دو جمله ای، هندسی، و پوسون ارائه کنید.

تمرین

- تمرین های ۹، ۲۰، ۲۱ و ۲۴، فصل چهارم کتاب شبیه سازی گسسته پیشامد
- تمرین های۵، ۹، ۱۹، و ۲۰ تا ۲۲ فصل هشتم کتاب شبیه سازی سیستم های گسسته پیشامد
- تمرین های۱۲، ۱۳، ۱۴، ۱۶، ۱۷، ۱۸ و ۲۱، فصل نهم کتاب شبیه سازی گسسته پیشامد
- چگونگی کاربرد رگرسیون خطی را در شبیه سازی توضیح دهید.
- با استفاده از روش همنهشتی خطی و پارامترهای زیر اعداد تصادفی مربوط را تولید کرده و طول دنباله تصادفی تولید شده را بدست آورید.

$a=4951$ $c=247$ $m=256$

Prepared By Dr. Kazemipoor