به نام خداوند بخشنده مهربان

روش‌های آماری و اقتصادسنجی

در تحلیل و مدل‌سازی داده‌های حمل و نقلی

محمد مهدی بشارتی

besharati@iut.ac.ir

توزیع های احتمالاتی

فصل چهارم

Probability distributions

مقدمه

بخش اول: توابع جرم احتمال و توزیع های گسسته

بخش دوم: توابع چگالی احتمال و توزیع های پیوسته

مقدمه

- احتمال وقوع (یا مشاهده) هریک از مقادیر ممکن برای یک متغیر تصادفی را چگونه برآورد کنیم؟
- برای مدل‌کردن مقادیر این متغیرهای تصادفی از تئوری احتمال استفاده می‌شود.
- مجموعه مقادیری که یک متغیر تصادفی می‌تواند به خود بگیرد، دارای یک توزیع احتمالاتی است.
- **توزیع احتمال**، نموداری‌است که احتمال وقوع هریک از مقادیر ممکن برای یک متغیرتصادفی را نمایش می‌دهد.
- رابطه‌ای که تمام مقادیر متغیرتصادفی(X) را به همراه احتمال‌های مربوطه نشان دهد؛ **تابع احتمال** $f(x)$ می‌نامند.

مقدمه

- اگر مقادیر ممکن (یا همان داده‌های مشاهده‌شده یا پیشامدها) به‌صورت اعداد صحیح باشند؛ متغیر مورد نظر، **متغیر گسسته** است. مثلا:

  - ✓ تعداد وسایل‌نقلیه واردشده به یک تقاطع در یک ساعت، تعداد عابرین‌پیاده، تعداد تصادفات سالانه در یک استان وغیره.

- از «**تابع جرم احتمال**(Probability Mass Function (PMF))»، برای برآورد احتمال مشاهده مقادیر متغیرهای **گسسته** استفاده می‌شود.

- برخی از توزیع‌های گسسته: ***توزیع دوجمله‌ای***، ***توزیع دوجمله‌ای منفی*** و ***توزیع پواسون***

مقدمه

- اگر مقادیر ممکن (یا همان داده‌های مشاهده‌شده یا پیشامدها) بصورت اعداد حقیقی باشند، متغیر مورد نظر، ***متغیر پیوسته*** است. مثلا:
  - ✓ مدت زمان سپری شده میان ورود دو خودرو متوالی به یک پارکینگ، زمان سفر در یک قطعه بزرگراهی، سرعت تردد خودروها در مقطع از یک جاده برون شهری وغیره.

- از «***تابع چگالی احتمال*** (Probability Density Function (PDF))»، برای برآورد احتمال مشاهده مقادیر متغیرهای **پیوسته** استفاده می‌شود.
- برخی از توزیع‌های پیوسته: ***توزیع‌های نرمال***، ***لگ‌نرمال***، ***نمایی***، ***گاما*** و ***کای‌دو***

توزیع های احتمالاتی

فصل چهارم

Probability distributions

مقدمه

بخش اول: توابع جرم احتمال و توزیع های گسسته

بخش دوم: توابع چگالی احتمال و توزیع های پیوسته

مدرس: محمدمهدی بشارتی

تابع جرم احتمال

- ✓ یک متغیر **گسسته**، دارای مقادیر معینی است که هرکدام از این مقادیر، با احتمال معینی، رخ می‌دهد.
- ✓ گاهی اوقات، احتمالات متناظر با هریک از این مقادیر را در جدول‌هایی به‌نام «**جدول فراوانی**» ارائه می کنند.

مقادیر محتمل	x	x_1	x_2	x_3	...	x_n
احتمال مشاهده هر مقدار	$f(x)$	$f(x_1)$	$f(x_2)$	$f(x_3)$	...	$f(x_n)$

تابع جرم احتمال

- ✓ اغلب آسان‌تر است که تمام احتمالات مربوط به یک متغیرتصادفی X را با فرمولی نمایش دهند که تابعی از مقادیر X بوده ($f(x)$) و به آن، **تابع احتمال** می گویند.
- ✓ مقدار این تابع به ازای $X=x$، احتمال رخ دادن پیشامدِ مشخص x را نشان می‌دهد.
- ✓ درحالتی‌که متغیر تصادفی گسسته باشد، تابع احتمال را ***تابع جرم احتمال*** می‌نامند.
- ✓ تابع جرم احتمال X، به‌ازای هرمقدار x در برد X به‌صورت زیر تعریف می‌شود:

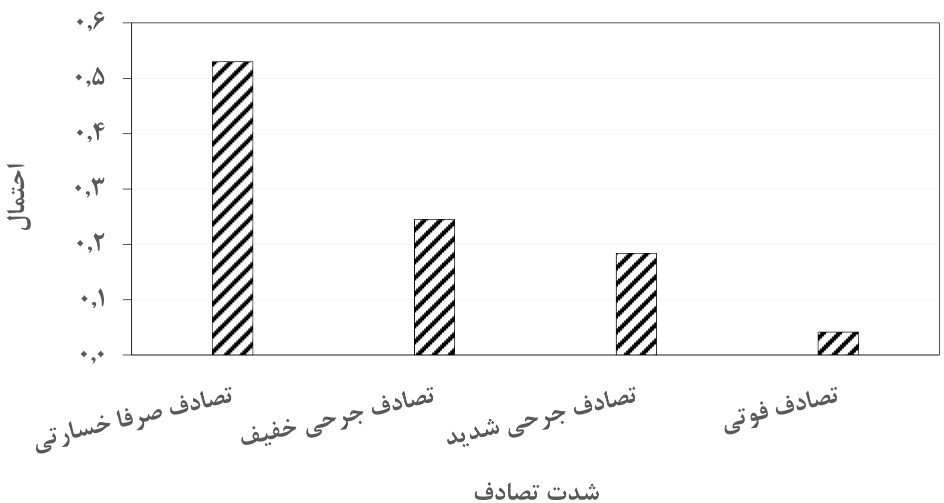
$$f(x) = P(X = x)$$

- ✓ همواره $\sum_i f(x_i) = 1$

توزیع احتمال گسسته

اگر مقادیر تابع احتمال یک متغیر تصادفی($f(x)$) را به ازای تمام xهای موجود در برد آن در یک نمودار به‌صورت زوج‌های ((x_i)،($f(x_i)$)) رسم کنیم؛ نمودار توزیع احتمال به‌دست می‌آید.

- به‌دلیل گسسته بودن مقادیر متغیر، نمودار توزیع احتمال یک متغیر گسسته به‌صورت یک **نمودار میله‌ای** و یا **مستطیلی** می‌باشد.
- هنگامی‌که فاصله بین xها مساوی باشد از نمودار مستطیلی و درغیراینصورت از نمودار میله‌ای استفاده می‌شود.



تابع توزیع(تابع توزیع تجمعی) گسسته

احتمال این که مقدار متغیرتصادفی گسسته X از عدد حقیقی x_0 کوچکتر یا مساوی باشد چقدر است؟

$$F(x) = P(X \leq x) = \sum_{z \leq x}^{n} f(z)$$

- پیش از توسعه نرم‌فزارهای آماری، مقـادیر تابع توزیع برای توزیع‌های آماری معروف در قالب جداول آماده ارائه می‌شد. اما امروزه از نرم افزارهای کامپیوتری استفاده می‌شود.
- **نمودار توزیع تجمعی احتمال:** مقادیر تابع توزیع تجمعی یک متغیر تصادفی($F(x)$) به ازای تمام xهای موجود در برد آن.
- نمودار توزیع تجمعی برای متغیرهای گسسته به‌صورت نمودار پله‌ای خواهد بود

آزمایش برنولی

سوال: آزمایش برنولی چیست؟ مثال از حوزه حمل ونقل و ترافیک؟

- آزمایش(یا مشاهده) برنولی، آزمایشی‌است که هربار انجام آن، به یکی از نتایج ناسازگار موفقیت یا شکست می‌انجامد.
- مثلا:
  - جنسیت هر مسافری که در هر لحظه وارد ایستگاه BRT می شود (مرد-زن)
  - سواد دار بودن راننده هر یک از اتوبوس‌های یک ترمینال (باسواد-بی سواد)
  - و ...

توزیع دو جمله ای

اگر یک آزمایش برنولی را n مرتبه به‌طور مستقل از هم تکرار کنیم (n) و احتمال موفقیت در هر آزمایش، ثابت باشد (p)؛ این نوع آزمایش را آزمایش دوجمله‌ای می‌گویند.

اگر متغیر تصادفی X را به صورت **«تعداد موفقیت‌ها در n آزمایش مستقل برنولی»** تعریف کنیم، تابع جرم احتمال آن به شرح زیر بوده و دارای توزیع دوجمله‌ای خواهد بود که با نماد Bin (n, p) نشان داده می‌شود؛

$$f(x) = P(X = x) = \binom{n}{x} p^x (1-p)^{n-x} \quad ، \; x=۰,۱,۲, \dots , n$$

X : (متغیرتصادفی) تعداد موفقیت‌ها در n آزمایش مستقل برنولی

n: تعداد دفعات انجام آزمایش‌های برنولی

p: احتمال موفقیت در هر آزمایش برنولی

چهار شرط کلاسیک توزیع دوجمله‌ای:

1. تعداد آزمایش‌ها ثابت است (n مشخص).
2. هر آزمایش فقط دو نتیجه دارد (موفقیت/عدم موفقیت).
3. آزمایش‌ها مستقل‌اند.
4. احتمال موفقیت در همه آزمایش‌ها برابر p است.

توزیع دو جمله ای

مثال:

یک آزمایش برنولی با احتمال موفقیت ۰/٤۵ را ۱۰ مرتبه تکرار کرده‌ایم.

مطلوبست محاسبه احتمال مشاهده ۰ ، ۱، ۲، ... ، ۱۰ پیروزی در این آزمایش دوجمله‌ای؟

$$f(4) = P(X = 4) = \binom{10}{4} 0.45^4 (1 - 0.45)^6$$

$$= \frac{10!}{4!\,6!} \times 0.041 \times 0.0277 = 0.2384$$

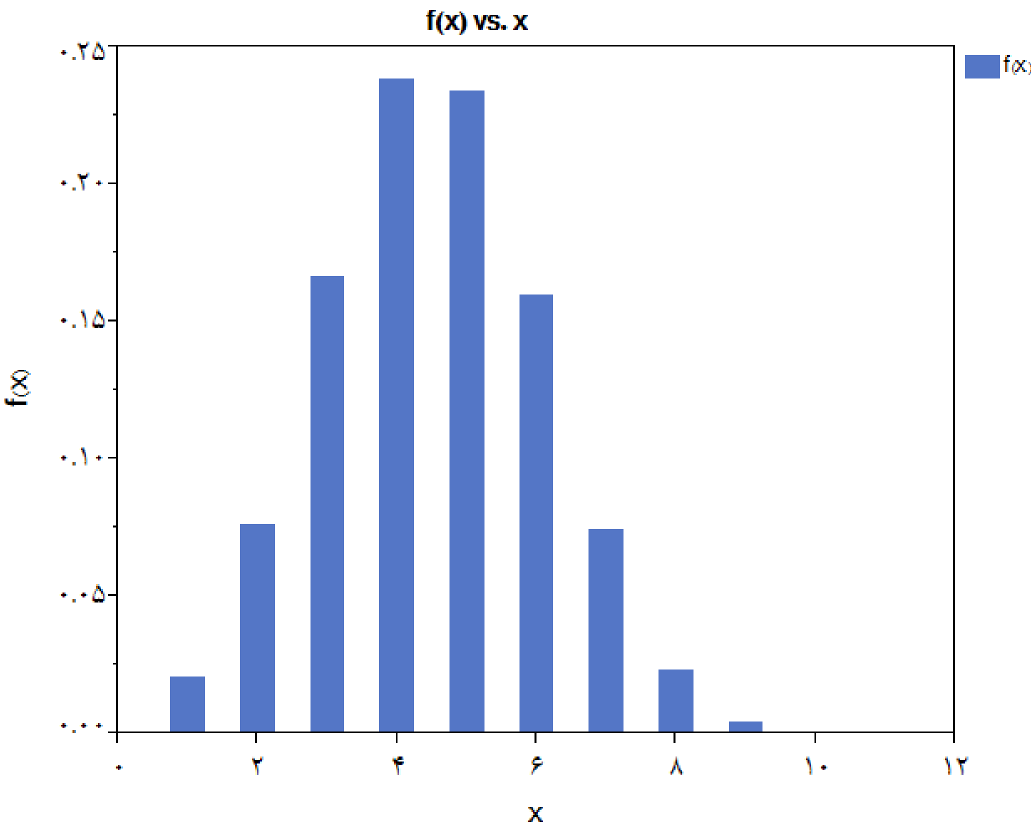
x	P(X=x)
1	0.0207
2	0.0763
3	0.1665
4	0.2384
5	0.2340
…	
10	0.0003

توابع جرم احتمال و توزیع های گسسته

توزیع دو جمله ای

در مثال قبل، اگر توزیع مقادیر احتمال به‌دست آمده برای مقادیر مختلف X را به‌صورت نمودار رسم کنیم، یک نمودار توزیع دوجمله‌ای خواهیم داشت.

x	P(X=x)
1	0.0207
2	0.0763
3	0.1665
4	0.2384
5	0.2340
…	
10	0.0003



توزیع دو جمله ای

امیدریاضی و واریانس توزیع دوجمله‌ای:

$$E(X) = np$$

$$\mathrm{Var}(X)=npq$$

توزیع دو جمله ای منفی

فرض کنید آزمایش‌های برنولی مستقل با احتمال موفقیت ثابت p انجام می‌دهیم.

- به‌جای اینکه بپرسیم: در n آزمایش چند موفقیت داریم؟ (مدل دوجمله‌ای)
- می‌پرسیم: چند آزمایش لازم است تا به r موفقیت برسیم؟ این دقیقاً مدل دوجمله‌ای منفی است.

توزیع دو جمله ای منفی

اگر آزمایش برنولی را آنقدر ادامه دهیم تا r موفقیت به‌دست آوریم و پس از مشاهده r–امین موفقیت، آزمایش متوقف شود؛ چنین آزمایشی را آزمایش دوجمله‌ای منفی می‌نامیم.

فرم اول تابع جرم احتمال متغیر تصادفی دوجمله‌ای منفی به‌صورت زیر است ($X \sim NB(r,p)$)؛

$$f(x)=P(X=x)=\binom{x-1}{r-1}\, p^{r}(1-p)^{x-r} \quad ، \; x=r, r+1, r+2, \ldots$$

r : تعداد موفقیت

X : تعداد دفعات انجام آزمایش تا رسیدن به r–امین موفقیت

p : احتمال موفقیت در هر آزمایش برنولی

توزیع دو جمله ای منفی

احتمال مشاهده r–امین موفقیت در x–امین آزمایش، در واقع برابر است با احتمال اینکه دو حادثه زیر رخ دهد؛

۱– (r-1) موفقیت در (x-1) آزمایش رخ دهد (احتمال دوجمله‌ای)،

۲– (r) امین موفقیت در (x) امین آزمایش رخ دهد (احتمال اینکه نتیجه آزمایش X–ام، موفقیت باشد)،

$$f(x)=P(X=x)=\binom{x-1}{r-1} p^r(1-p)^{x-r} \quad ، x=r, r+1, r+2, \ldots$$

$$E(X)=\frac{r}{p} \qquad Var(X)=\frac{r(1-p)}{p^2}$$

توزیع دو جمله ای منفی

اگر آزمایش برنولی را آنقدر ادامه دهیم تا r موفقیت به‌دست آوریم و پس از مشاهده r–امین موفقیت، آزمایش متوقف شود؛ چنین آزمایشی را آزمایش دوجمله‌ای منفی می‌نامیم.

فرم دوم تابع جرم احتمال متغیر تصادفی دوجمله‌ای منفی به‌صورت زیر است ($Y \sim NB(r,p)$)؛

$$f(y)=P(Y=y)=\binom{y+r-1}{r-1}\, p^{r}(1-p)^{y} \quad ، \; y=0, 1, 2, \ldots$$

r : تعداد موفقیت

Y : تعداد شکست‌ها تا قبل از رسیدن به r–امین موفقیت

p : احتمال موفقیت در هر آزمایش برنولی

$$E(Y)=\frac{r(1-p)}{p} \qquad Var(X)=\frac{r(1-p)}{p^2}$$

توزیع دو جمله ای منفی

چند نکته؛

$X=Y+r$

فرم اول کل آزمایش‌ها را شمارش می‌کند (رایج در کتب تئوری احتمال)

فرم دوم فقط شکست‌ها را شمارش می‌کند (رایج برای مدلسازی داده‌های شمارشی و لحاظ نمودن بیش‌پراکندگی).

- ✓ در مدلسازی رگرسیون دوجمله‌ای منفی (پاورپوینت ۱۲) از ایده فرم دوم استفاده می‌شود.

توزیع دو جمله ای منفی

❖ مثال

از یک تکنسین خواسته‌شده که هرچه سریع‌تر ۳ تابلو هشداردهنده نصب‌شده در حاشیه یک جاده را با تابلوهای سالم تعویض کند. بر اساس تجربه گذشته خود، این تکنسین انتظار دارد که ٪۷۰ از تابلوهای موجود در انبار برای نصب سریع، آماده بوده و ٪۳۰ مابقی برای استفاده نیازمند تغییر و تعمیر است. درصورتی‌که در حال حاضر ۵ تابلو در انبار وجود داشته باشد، چقدر احتمال دارد که تکنسین، سومین تابلو موردنیاز خود را در هنگام بررسی پنجمین تابلو بیابد؟

r = تعداد پیروزی (تعداد تابلوهای آماده برای نصب سریع)= ۳

x = تعداد کل مشاهدات (تعداد تابلوهای موجود در انبار)= ۵

p= احتمال پیروزی (مشاهده تابلوی آماده برای نصب)= ۰/۷

$$P(X=۵)=\frac{(x-۱)!}{(r-۱)!\times(x-r)!}\times p^r(۱-p)^{x-r}=\frac{(۵-۱)!}{(۳-۱)!\times(۵-۳)!}\times(۰/۷)^۳(۰/۳)^۲=۰/۱۸۵$$

بنابراین، احتمال اینکه پس از مشاهده پنجمین تابلو درون انبار، مجموعه ۳ تابلو موردنیاز یافته شود برابر با ۰/۱۸۵ می‌باشد. توجه شود که این احتمال که پس از مشاهده سومین تابلو و یا چهارمین تابلو درون انبار، سومین تابلو سالم موردنیاز یافته شود نیز وجود دارد. به‌عنوان تمرین، این احتمالات را محاسبه نمایید.

توزیع هندسی

اگر در آزمایش دوجمله‌ای منفی، تعداد آزمایش‌ها را تا رسیدن به **اولین موفقیت (شکست)** ادامه دهیم؛ به چنین آزمایشی، آزمایش هندسی گفته می‌شود. تابع توزیع هندسی به صورت زیر است؛

$$f(x)=P(X=x)=p(1-p)^{x-1},\ x=1, 2, 3, \ldots$$

X : تعداد دفعات انجام آزمایش تا مشاهده اولین موفقیت (شکست)
p : احتمال موفقیت (شکست) در هر آزمایش برنولی

✓ پس توزیع هندسی حالت خاصی از دوجمله‌ای منفی است وقتی r=1

مثال‌هایی از رویدادهای دارای توزیع هندسی؟

- بازرسی از محصولات کارخانه تا زمانی ادامه می‌یابد تا اولین محصول خراب مشاهده شود.
- پیچ‌های گاردریل‌ها بازدید شود تا زمانی که اولین پیچ خراب مشاهده شود.
- ....؟

توزیع پواسون

اگر متغیر تصادفی X بصورت «**تعداد دفعات وقوع یک رخداد در یک بازه زمانی یا فاصله مکانی معین**» تعریف شود (و در این شمارش شرایط فرآیند پواسون برقرار باشد)، آنگاه می‌توان گفت متغیر X از توزیع پواسون پیروی می‌کند. برای مثال:

- تعداد تصادفات ترافیکی سالانه در یک تقاطع،
- تعداد وسایل‌نقلیه ورودی به یک تقاطع در یک بازه زمانی معین
- تعداد عابرین پیاده که در یک ساعت به یک تقاطع می‌رسند
- **و ....**

در موارد بالا، می‌توان فراوانی وقوع رویداد موردنظر در یک بازه‌زمانی معین را به کمک توزیع پواسون مدل کرد.

احتمال اینکه یک حادثه به تعداد مشخصی در فاصله زمانی یا مکانی ثابتی رخ دهد را شرح می‌دهد؛ به شرط اینکه این حوادث با نرخ وقوع مشخص (ثابت) و مستقل از زمان آخرین حادثه رخ دهند.

توزیع پواسون

یک آزمایش پواسون، **لزوماً** دارای خصوصیات زیر است:

- میانگین تعداد موفقیت‌ها در بازه زمانی یا مکانی مشخص‌شده (نرخ وقوع)، **ثابت و معلوم** است.
- تعداد موفقیت‌های مشاهده شده در یک بازه زمانی (مکانی) معین، **مستقل** از تعداد موفقیت‌های مشاهده شده در یک بازه زمانی (مکانی) مشخص دیگر است.
- احتمال وقوع **بیش از یک موفقیت** در یک بازه زمانی کوتاه (ناحیه مکانی کوچک)، **قابل اغماض** باشد.
- احتمال موفقیت در هر بازه زمانی (ناحیه مکانی)، متناسب با آن فاصله زمانی (مکانی) بوده و به تعداد موفقیت‌هایی که در خارج از آن فاصله زمانی (مکانی) رخ می‌دهد، بستگی نداشته باشد.

توزیع پواسون

تابع جرم احتمال توزیع پواسون به صورت زیر است

$$f(X) = P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!} \qquad , \qquad x=۰, ۱, ۲, \ldots$$

X : تعداد دفعات وقوع رخداد موردنظر در فاصله زمانی (مکانی) معین

λ: میانگین تعداد دفعات وقوع رخداد موردنظر در فاصله زمانی (مکانی) معین

یکی از ویژگی‌های توزیع پواسون آنست که **میانگین و واریانس این توزیع با هم برابر بوده و مساوی λ** می‌باشد.

$$E(X) = \lambda \qquad Var(X) = \lambda$$

- ✓ توزیع پواسون می‌تواند برای مدل‌سازی رویدادهای بکار برده شود که دارای تعداد وقایع بالقوه بسیار زیاد هستند و احتمال وقوع هر واقعه بسیار کم است.

یکی از مهم‌ترین کاربردهای این توزیع در مهندسی حمل‌ونقل، مدلسازی فراوانی تصادفات می‌باشد.
مقادیر $P(X \le r)$ برای توزیع پواسون به ازای مقادیر مختلف λ در جداول آماری ارائه می‌شود.

توزیع پواسون

مثال:

میانگین تعداد کشتی‌هایی که روزانه وارد یک بندر می‌شوند، ۱۰ فروند است. زیرساخت‌های این بندر قادر است حداکثر به ۱۵ فروند کشتی در روز سرویس دهد. مطلوبست محاسبه احتمال اینکه در یک روز معین، این بندر نتواند پاسخگوی تعداد کشتی‌هایی باشد که در آن پهلو می‌گیرند؟

اگر تعداد کشتی‌های وارد شده به بندر را بعنوان متغیر تصادفی X درنظر بگیریم؛ می‌توان گفت در این مسئله، X از توزیع پواسون با $\lambda=10$ پیروی می‌کند. بدین‌ترتیب،

$$P(X>15)=1-P(X\leq 15)=1-\sum_{x=0}^{15}p(x;10)=1-0.9513\ =\ 0.0487$$

با توجه به جدول توزیع پواسون، $\sum_{x=0}^{15}p(x;10)=0.9513$ می‌باشد.

توزیع پواسون

نکته:

اگر تعداد دفعات وقوع رخداد موردنظر در واحد زمان از توزیع پواسون با پارامتر λ پیروی کند، آنگاه، تعداد دفعات وقوع رخداد موردنظر در t واحد زمانی از توزیع پواسون با پارامتر $\lambda.t$ پیروی خواهد کرد. درواقع، در این حالت، X، تعداد دفعات وقوع رخداد موردنظر در t واحدزمانی خواهد بود.

$$f(X)=P(X=x)=\frac{e^{-\lambda t}\,\lambda t^{x}}{x!}$$

تقریب توزیع دوجمله ای با توزیع پواسون

از توزیع پواسون اغلب برای مدل‌سازی رویدادهایی استفاده می‌شود که دارای تعداد وقایع بالقوه بسیار زیاد هستند؛ اما احتمال وقوع هر واقعه بسیار کم است.

هرگاه در توزیع دوجمله‌ای، n به سمت بینهایت و p به سمت صفر میل کند، آنگاه می‌توان فرض کرد که میانگین موفقیت‌ها $(n.p)$ ثابت بوده و برابر با λ است. در این حالت می‌توان احتمال‌های متناظر با **توزیع دوجمله‌ای** را با **توزیع پواسون** تقریب زد

$$X_n \sim Binomial\ (n, p)$$

و n به سمت بینهایت و p به سمت صفر میل کند. به گونه‌ای که: $np \to \lambda$

$$\mathrm{Bin}(n,p) \approx Poisson(\lambda = np)$$

✓ برای nهای بزرگ (بزرگتر از ۲۰) و pهای کوچک (کوچکتر از ۰.۰۵) می‌توان از این تقریب استفاده کرد.

تقریب توزیع دوجمله ای با توزیع پواسون

مثال:

براساس مطالعات گذشته بر روی جاده‌های یک استان، مشخص شده که به‌طور متوسط در ۱٪ موارد، پایه گاردریل به ریل متصل نشده‌است. مطلوبست احتمال اینکه با بررسی تصادفی ۲۰۰ پایه گاردریل در راه‌های این استان، هیچ پایه‌ی بدون اتصال به گاردریل مشاهده نشود؟

پاسخ: بررسی اتصال یا عدم اتصال هر پایه به گاردریل، به منزله یک آزمایش برنولی بوده که از قبل می‌دانیم احتمال موفقیت در هریک از این آزمایش‌ها (یعنی مشاهده عدم‌اتصال پایه به گاردریل) برابر با ۰/۰۱ است. اگر متغیر تصادفی X را بصورت تعداد پایه‌های بدون‌اتصال در میان ۲۰۰ پایه مشاهده شده درنظر بگیریم؛ آنگاه می‌توان گفت X در این مسئله از توزیع دوجمله‌ای پیروی می‌کند. بنابراین

$$P(X=0)=\binom{200}{0}0.01^0(0.99)^{200-0}=0.134$$

حل با استفاده از تقریب توزیع پواسون. از هر ۱۰۰ پایه، یک پایه بدون‌اتصال وجود دارد($\lambda=1$). به‌عبارت دیگر، می‌توان گفت انتظار می‌رود بطور متوسط، به‌ازای هر ۲۰۰ پایه، ۲ عدد بدون‌اتصال باشد. بنابراین

$$P(X=0)\cong\frac{e^{-2}.2^0}{0!}=0.1353$$

توزیع فوق هندسی

مجموعه‌ای از N عضو **متناهی** را در نظر بگیرید که K عضو آن دارای یک ویژگی و بقیه، فاقد این ویژگی هستند.

مانند ۵۰۰ لامپ موجود در یک جعبه که ۳۰۰ تای آن سالم و بقیه معیوب باشند.

حال فرض کنید می خواهیم از این مجموعه، n عضو به صورت تصادفی (**بدون جایگذاری**) انتخاب کنیم.

دراین صورت اگر متغیر تصادفی X تعداد عناصری در n برداشت باشد که دارای ویژگی موردنظر هستند، آنگاه X دارای توزیع فوق هندسی است. K,N و n اعداد صحیح و مثبت‌اند.

N: تعداد اعضای مجموعه

K: تعداد اعضای جامعه که دارای ویژگی خاص هستند.

n: تعداد اعضای انتخاب شده از درون مجموعه

k: تعداد اعضای نمونه که دارای ویژگی خاص هستند

متغیر تصادفی X: تعداد اعضای دارای ویژگی خاص از میان n عضو انتخاب شده

تابع جرم‌احتمال توزیع فوق‌هندسی

$$p(k) = p(X = k) = \frac{\binom{K}{k}\binom{N-K}{n-k}}{\binom{N}{n}}$$

توجه: **توزیع دوجمله‌ای** احتمال k موفقیت در n مشاهده را **با جایگذاری** محاسبه می‌کند.

تقریب توزیع فوق هندسی با توزیع دوجمله‌ای

هنگامی که اندازه جامعه N نسبت به اندازه نمونه n خیلی بزرگتر باشد؛ آنگاه نمونه‌گیری بدون جایگذاری تفاوت چندانی با نمونه‌گیری با جایگذاری نخواهد داشت.

در این شرایط می‌توان از توزیع دوجمله‌ای به جای توزیع فوق هندسی استفاده کرد.

اهمیت این موضوع آن است که توزیع دوجمله‌ای پارامترهای کمتری دارد و نیاز به دانستن اندازه جامعه و نیز تعداد اعضای دارای ویژگی خاص در جامعه نیست؛ و صرفاً کافیست که نسبت آن‌ها را داشته باشیم.

به طور کلی با افزایش N به سمت بینهایت، برای یک n ثابت، توزیع فوق هندسی با پارامترهای N، K، و n به سمت توزیع دوجمله‌ای با پارامترهای n و p میل می‌کند (p همان نسبت K به N (احتمال موفقیت) است).

$$p = \frac{K}{N}$$

$$\mathrm{Bin}(n,p) \approx Hypergeometric(N,K,n)$$

توزیع های احتمالاتی

فصل چهارم

Probability distributions

مقدمه

بخش اول: توابع جرم احتمال و توزیع های گسسته

بخش دوم: توابع چگالی احتمال و توزیع های پیوسته

توابع چگالی احتمال و توزیع های پیوسته

مدرس: محمدمهدی بشارتی

تابع چگالی احتمال،

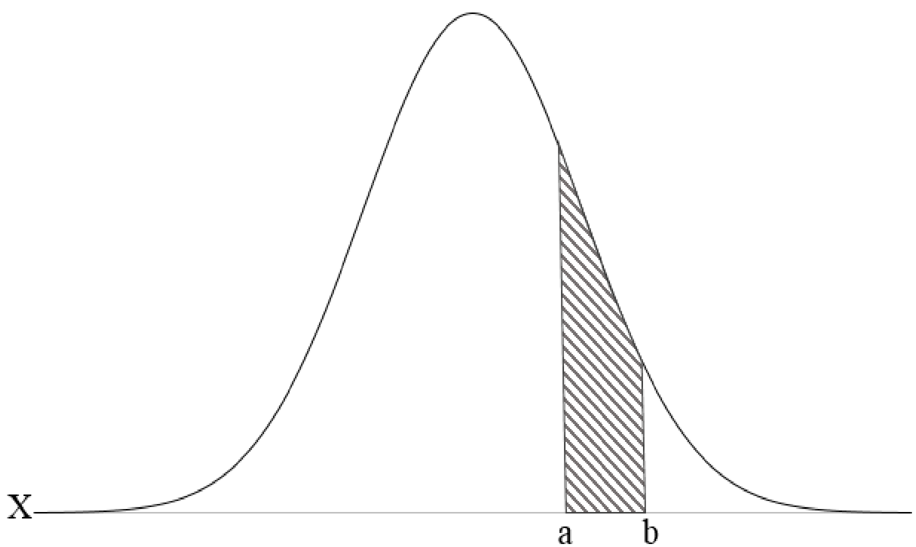
- هنگامی که متغیر تصادفی پیوسته باشد، تابع احتمال را تابع چگالی احتمال می‌نامند ($f(x)$).
- از آنجا که یک متغیر پیوسته دارای بینهایت مقدار است، احتمال اینکه متغیرتصادفی مذکور، دقیقا با یکی از این بینهایت مقدار، برابر شود، مساوی با صفر است.
- به‌همین دلیل، نمی‌توان تابع چگالی احتمال یک متغیرتصادفی پیوسته را در قالب جدول ارائه داد؛ بلکه چنین توابعی، تنها به‌صورت فرمول بیان می‌شوند.
- سوال: تابع توزیع را می‌توان به صورت جدول ارائه کرد؟

تابع چگالی احتمال،

- ✓ مقدار انتگرال تابع چگالی احتمال($f(x)$)، برابر است با احتمال اینکه متغیر X، مقداری بین a و b داشته باشد

$$P(a \leq X \leq b) = \int_a^b f(x)dx$$

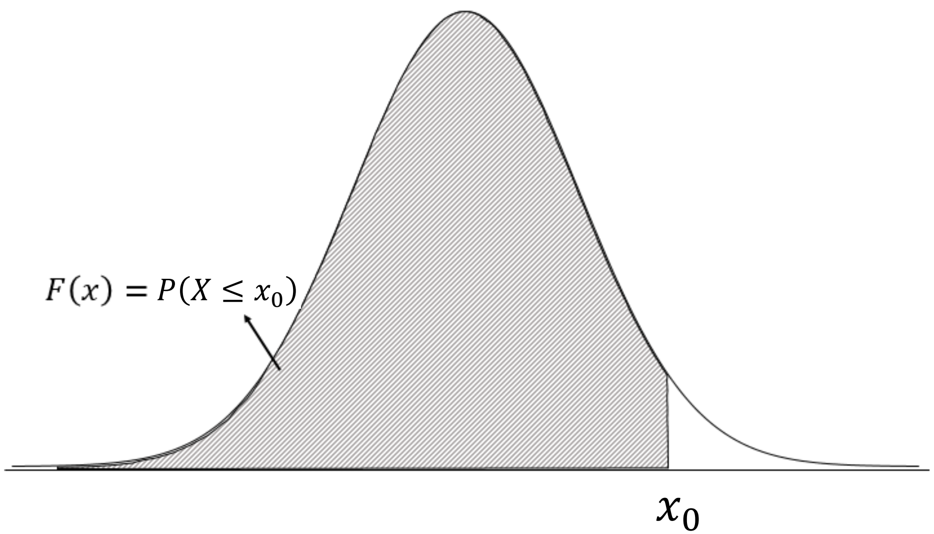
- ✓ در این حالت، احتمال اینکه پیشامد به‌خصوص x_0 رخ دهد، برابر با سطح زیر منحنی توزیع احتمال خواهد بود. به همین دلیل رخدادها به‌صورت بازه‌ای از پیشامدها نمایش داده می‌شوند.



تابع چگالی احتمال،

- ✓ برای مثال، اگرمتغیر تصادفی X را به‌صورت سرعت وسایل‌نقلیه در یک کمان درنظر بگیریم؛ آنگاه، احتمال اینکه یک وسیله نقلیه از سرعت مجاز x_0 تجاوز کند، برابر با سطح زیر منحنی توزیع احتمال برای تمامی مقادیر $x > x_0$ خواهد بود. این مقدار بصورت زیر محاسبه می‌شود؛

$$P(X > x_0) = 1 - P(X \le x_0)$$



تابع چگالی احتمال،

سوال: احتمال اینکه سرعت یک خودرو دقیقا با سرعت ۷۰ کیلومتر بر ساعت باشد، چقدر خواهد بود؟

پاسخ:

احتمال اینکه سرعت یک خودرو دقیقا با سرعت ۷۰ کیلومتر بر ساعت باشد قابل محاسبه نخواهد بود و این احتمال برابر با صفر بدست می‌آید.

هرچند به‌جای محاسبه این احتمال، می‌توان احتمال اینکه یک خودرو با سرعت حدودا ۷۰ کیلومتربرساعت حرکت کند را به‌دست آورد.

برای این منظور می‌توان احتمال اینکه x بین ٦٩ و ۷۱ باشد را محاسبه نمود($P(69 < X < 71)$).

تابع توزیع احتمال،

برای یک متغیر پیوسته، توزیع احتمال به‌صورت یک منحنی پیوسته خواهد بود که مقادیر احتمال در محور قائم و مقادیر متغیر تصادفی در محور افقی قرار می‌گیرد.

همچون حالت گسسته، مسائل زیادی وجود دارند که در آنها تحلیل‌گر تلاش می‌کند احتمال این که مقدار متغیرتصادفی پیوسته از عدد حقیقی x_0 کوچکتر یا مساوی باشد را تعیین نماید. در این راستا، مفهوم تابع توزیع (یا تابع توزیع تجمعی) را به‌صورت زیر تعریف می‌کنیم:

اگر X متغیرتصادفی پیوسته‌ای باشد که مقدار چگالی احتمال آن به‌ازای z برابر با $f(z)$ باشد، آنگاه تابع توزیع احتمال تجمعی به‌صورت زیر تعریف می‌گردد.

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(z)dz$$

تابع توزیع احتمال،

برای بسیاری از توزیع‌های معروف آماری، مقدار تابع توزیع در جداول آماده ارائه می‌شود.

هرچند با توسعه نرم‌افزارهای آماری، این جداول آماده کمتر مورد استفاده قرار می‌گیرند.

از میان توابع چگالی احتمال،

- برخی (مانند لگ نرمال، نمایی) غالبا برای مدل‌سازی رویدادهای فیزیکی،
- برخی دیگر (مانند توزیع t، توزیع F، توزیع کای‌دو) غالبا برای آزمون‌های آماری
- بعضی نیز برای هردو منظور (توزیع نرمال، گاما، بتا)

مورد استفاده قرار می‌گیرند.

توزیع یکنواخت،

ساده‌ترین نوع از توزیع‌های پیوسته می‌باشد.

متغیر X با تابع چگالی زیر، در بازه [a,b] دارای توزیع یکنواخت است؛

$$f(x) = \begin{cases} \dfrac{1}{b-a} & a \le x \le b \\ 0 & \text{در سایر نقاط} \end{cases}$$

هرگاه احتمال قرار گرفتن یک نقطه در قسمتی از این بازه، متناسب با طول آن بازه بوده و قرار گرفتن آن نقطه در درون آن بازه، یک پیشامد حتمی باشد، آنگاه می‌گوییم متغیر تصادفی X در بازه [a,b] به‌طور یکنواخت توزیع شده‌است.

توزیع یکنواخت،

مثال:

در یک ایستگاه اتوبوس معین، اتوبوس‌ها به فاصله ۲۰ دقیقه از یکدیگر حرکت می‌کنند(مثلا اگر اولی ساعت ٦ حرکت می‌کند، بعدی ساعت ٦ و ۲۰ دقیقه و ...). اگر زمان ورود یک مسافر به این ایستگاه، به صورت تصادفی و با توزیع یکنواخت بین ساعت ۷:۲۰ تا ۸ باشد، مطلوبست محاسبه احتمال این که این مسافر؛

الف) کمتر از ۱۰ دقیقه منتظر حرکت اتوبوس مانده باشد.

ب) بیشتر از ۱۵ دقیقه منتظر حرکت اتوبوس مانده باشد.

توزیع یکنواخت،

مثال:

اگر متغیر X را برابر با زمان ورود این مسافر درنظر بگیریم، چون X دارای توزیع یکنواخت است، بنابراین، تابع چگالی احتمال زمان انتظار این مسافر بصورت زیر خواهد بود:

$$f(x) = \frac{1}{8 - 7:20} = \frac{1}{40}$$

الف) شرط لازم و کافی برای اینکه این مسافر کمتر از 10 دقیقه منتظر حرکت اتوبوس مانده باشد، برابر است با احتمال اینکه این مسافر در بازه‌های زمانی *7:30* تا *7:40* و یا *7:50* تا *8* وارد ایستگاه شده‌باشد.

$$P(7:30 < x < 7:40) + P(7:50 < x < 8) = \int_{30}^{40} \frac{1}{40} dx + \int_{50}^{60} \frac{1}{40} dx = 0.5$$

توزیع یکنواخت،

مثال:

ب) اگر این مسافر در بازه‌های زمانی *7:20* تا *7:25* و یا *7:40* تا *7:45* وارد ایستگاه شده‌باشد؛ مجبور است بیش از ۱۵ دقیقه منتظر حرکت اتوبوس بماند.

$$P(7{:}20 < x < 7{:}25) + P(7{:}40 < x < 7{:}45) = \int_{20}^{25} \frac{1}{40} dx + \int_{40}^{45} \frac{1}{40} dx = 0.25$$

توزیع نرمال،

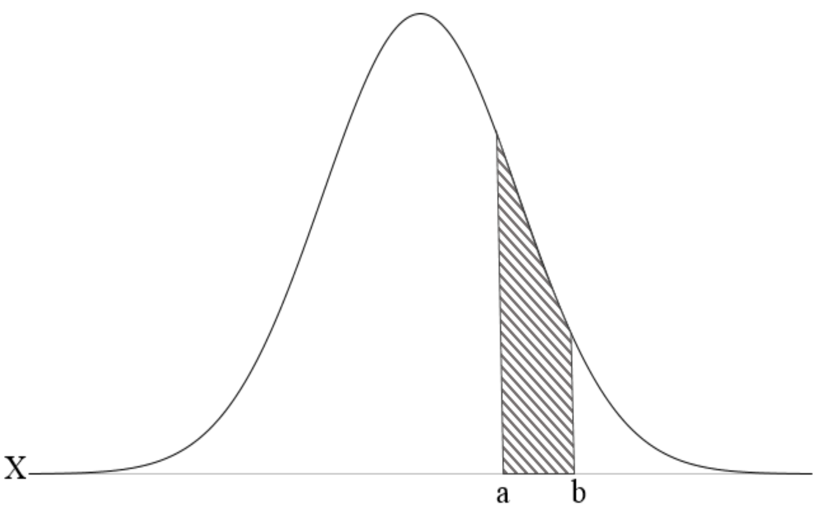
- پرکاربردترین توزیع پیوسته در مهندسی حمل‌ونقل
- نام های دیگر: توزیع گاوسی، نمودار زنگوله‌ای
- یک متغیر تصادفی با امیدریاضی μ و واریانس σ^2 دارای توزیع نرمال است؛ اگر تابع چگالی احتمال آن به شرح زیر باشد؛

$$f(x;\mu,\sigma^2)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}}\qquad -\infty\le x\le+\infty$$

توابع چگالی احتمال و توزیع های پیوسته

توزیع نرمال،

- عموماً توزیع نرمال به صورت $N(\mu,\sigma^2)$ نمایش داده می‌شود.
- منحنی نرمال نسبت به میانگین متقارن بوده و میانه، مد و میانگین آن بر هم منطبق هستند
- با انتگرال‌گیری از تابع چگالی توزیع نرمال، به‌ازای یک بازه معین از x (مثلا a تا b) و با داشتن مقادیر میانگین و واریانس توزیع، می‌توان مقادیر احتمال مشاهده بازه موردنظر را به‌دست آورد.



تابع چگالی نرمال استاندارد

به ازای مقادیر مختلف و ، شکل توزیع نرمال متفاوت خواهد بود.

توزیع نرمال استاندارد، حالت خاصی از توزیع نرمال است که در آن، متغیر تصادفی Z را به صورت زیر تعریف می‌کنیم؛

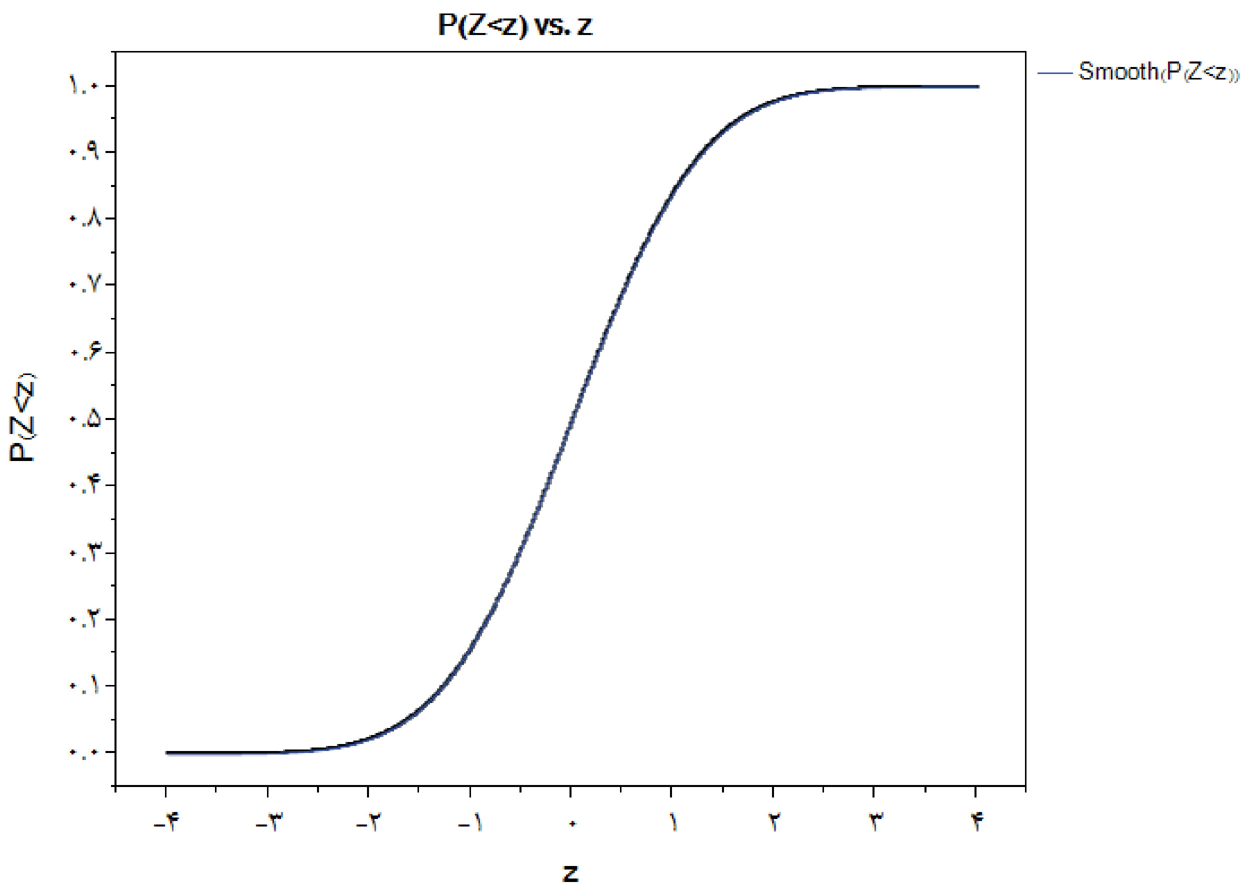
$$z = \frac{x - \mu}{\sigma}$$

با انتگرال گیری از تابع چگالی احتمال

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$$

- ✓ به ازای بازه‌های مختلف Z، می‌توان احتمالات متناظر با مشاهده این مقادیر را بدست آورد.
- ✓ مقادیر $P(Z < z)$ به ازای مقادیر مختلف Z محاسبه شده و در جدول‌های آماده، ارائه می‌شود.

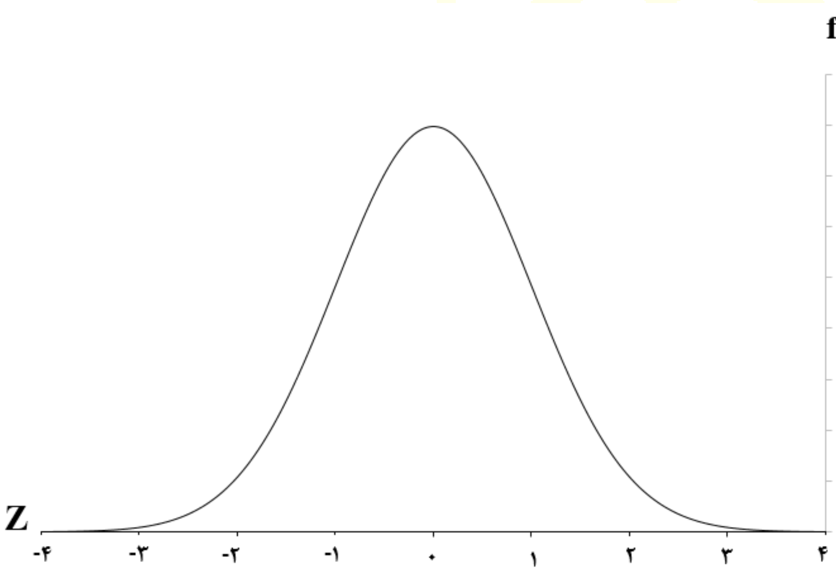
تابع توزیع نرمال استاندارد (توزیع تجمعی)



مقادیر $P(Z < z)$

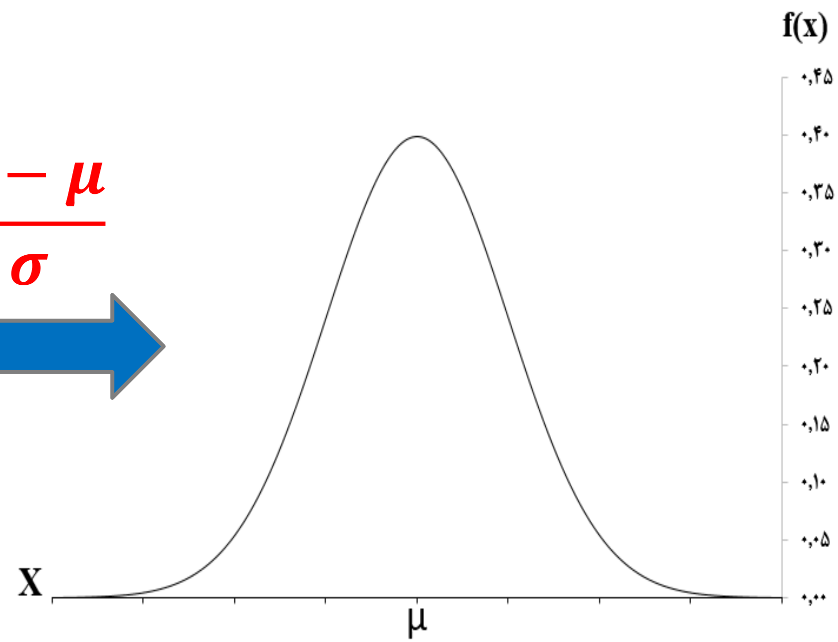
توزیع نرمال استاندارد (توزیع چگالی)

توزیع نرمال با $\mu = 0$ و $\sigma = 1$ به عنوان توزیع نرمال استاندارد شناخته شده و با نماد $N(0,1)$ نمایش داده می‌شود.



توزیع نرمال استاندارد

$$z = \frac{x-\mu}{\sigma}$$



توزیع نرمال

توزیع نرمال استاندارد (توزیع چگالی)

مثال:

فرض کنید متغیر تصادفی X دارای توزیع نرمال با میانگین ۱۰۰ و واریانس ٤ باشد. احتمال اینکه متغیر تصادفی X کمتر از ۱۰۲/۷ باشد، چقدر است؟

ابتدا متغیر تصادفی X را به متغیر تصادفی Z تبدیل کرده و سپس از جداول آماده استفاده می‌کنیم؛

$$z=\frac{x-\mu}{\sigma} \quad \Longrightarrow \quad z=\frac{x-100}{2}$$

$$z=\frac{102.7-100}{2}=1.35 \qquad P(Z\le 1.35)=0.9115$$

توزیع نرمال استاندارد (توزیع چگالی)

یکی از کاربردهای توزیع‌های آماری، تعیین عددی است که درصد معینی از داده‌ها، مقداری کمتر از آن داشته‌باشند.

با توجه به اینکه

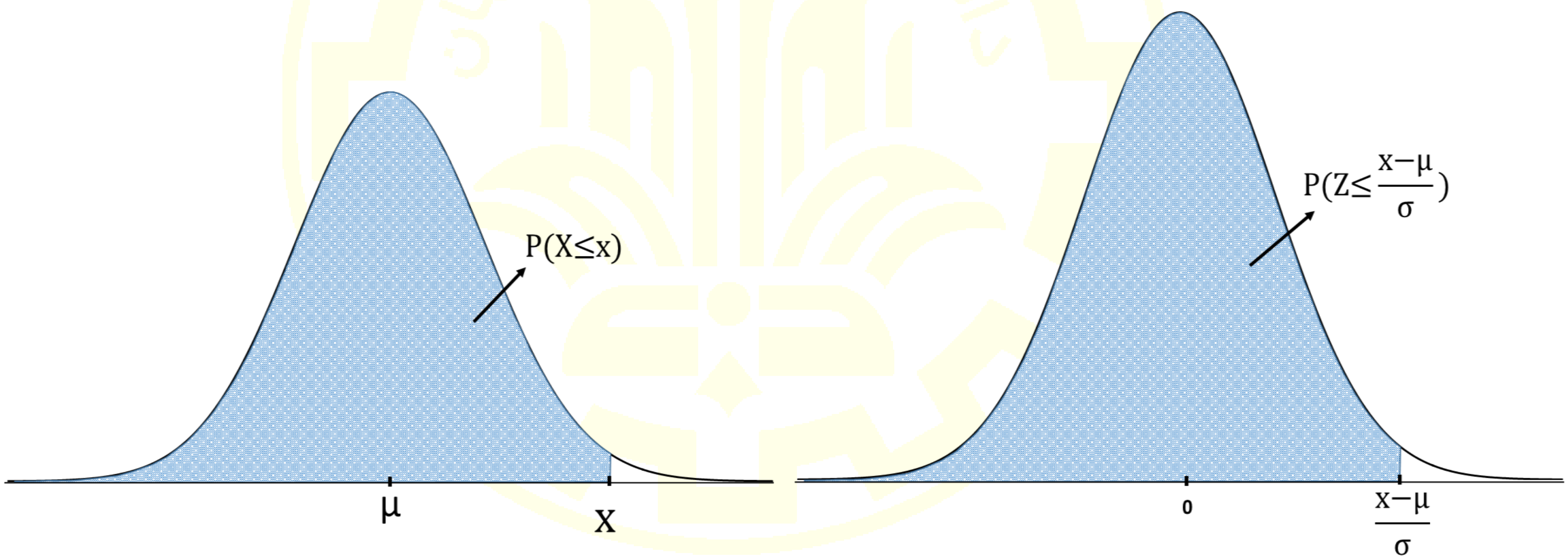
$$z = \frac{x-\mu}{\sigma} \Rightarrow x = \mu + Z * \sigma$$

با داشتن مقادیر **صدک‌های مربوط به توزیع نرمال استاندارد** و قرار دادن مقادیر مورد نظر در معادله $x = \mu + Z * \sigma$ می‌توان مقادیر صدک‌های مربوط به هر توزیع نرمال دیگری را به‌دست آورد.

صدک‌های توزیع نرمال استاندارد و مقادیر احتمال متناظر با آنها در جداول آماده، تهیه شده و مورد استفاده تحلیل‌گران قرار می‌گیرد.

توزیع نرمال استاندارد (توزیع چگالی)

صدک‌های توزیع نرمال استاندارد و مقادیر احتمال متناظر با آنها در جداول آماده، تهیه شده و مورد استفاده تحلیل‌گران قرار می‌گیرد.



توزیع نرمال استاندارد (توزیع چگالی)

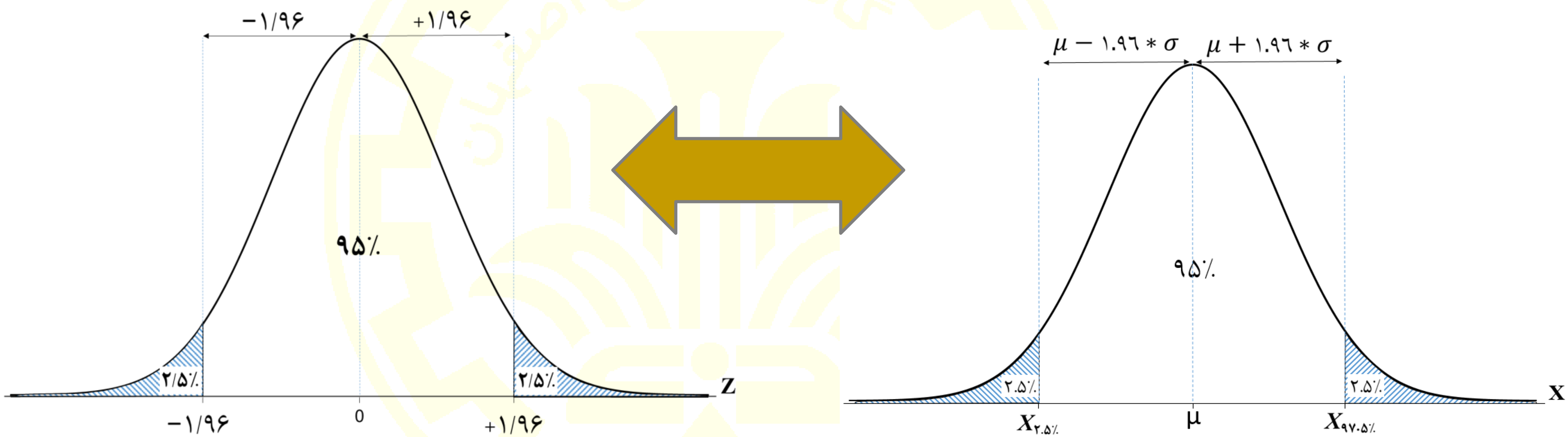
برای مثال، زمانی‌که گفته می‌شود مقدار 97.5 امین درصد داده‌ها در توزیع نرمال استاندارد حدودا برابر با 1.96 است بدین معناست که در هر توزیع نرمالی، 95٪ از (مقادیر) داده‌ها در بازه بین $\pm 1.96\sigma$ از میانگین، قرار می‌گیرد. برای مثال،

- **68.3** ٪ از مشاهدات در بازه $\mu \pm 1 * \sigma$ ،
- **95** ٪ از مشاهدات در بازه $\mu \pm 1.96 * \sigma$ ،
- **99.7** ٪ از مشاهدات در بازه $\mu \pm 3 * \sigma$

قرار خواهد داشت. این موضوع را می‌توان به شکل دیگری نیز تفسیر نمود. احتمال اینکه مشاهده بعدی در بازه $(\mu \pm Z \times \sigma)$ باشد، چقدر خواهد بود.

توزیع نرمال و نرمال استاندارد

احتمال اینکه مشاهده بعدی در بازه $(\mu \pm Z \times \sigma)$ باشد، چقدر خواهد بود.



ب) بازه ۹۵٪ مرکزی توزیع نرمال استاندارد

الف) بازه ۹۵٪ مرکزی یک توزیع نرمال

توزیع نرمال و نرمال استاندارد

مثال:

اگر متغیر تصادفی X دارای توزیع نرمال با میانگین μ و انحراف استانداردσ باشد، مطلوبست:

$$P(\mu - 1 \times \sigma < X < \mu + 1 \times \sigma)$$

$$P(\mu - 2 \times \sigma < X < \mu + 2 \times \sigma)$$

$$P(\mu - 3 \times \sigma < X < \mu + 3 \times \sigma)$$

پاسخ.

$$P(\mu - 1 \times \sigma < X < \mu + 1 \times \sigma) = P(-1 < Z < +1) = 0.8413 - 0.1587 = 0.6828$$

$$P(\mu - 2 \times \sigma < X < \mu + 2 \times \sigma) = P(-2 < Z < +2) = 0.9772 - 0.0228 = 0.9544$$

$$P(\mu - 3 \times \sigma < X < \mu + 3 \times \sigma) = P(-3 < Z < +3) = 0.9987 - 0.0013 = 0.9974$$

توزیع نرمال و نرمال استاندارد

مثال:

متغیر X دارای توزیع نرمال با میانگین ۵۵ و انحراف معیار ۳ می‌باشد. مطلوبست تعیین بازه‌ای که ۲.۵٪ از داده‌ها از آن کمتر و بیشتر باشند(۹۵٪ میانی)؟

پاسخ:

در هر توزیع نرمال، 95 ٪ از مشاهدات در بازه $\mu \pm 1.96 * \sigma$ است.

۲.۵ امین صدک برای این توزیع برابر است با: ۵۵+۳×(−۱.۹۶) = ۵۵−۳×(۱.۹۶) = ۴۹.۱۲

۹۷.۵ امین صدک برابر است با: ۵۵+۳×(۱.۹۶)=۶۰.۸۸

بدین‌ترتیب، می‌توان گفت که ۹۵٪ مرکزی داده‌ها، مقداری بین ۴۹.۱۲ و ۶۰.۸۸ دارند.

تقریب توزیع دوجمله‌ای با توزیع نرمال

اگر X یک متغیر تصادفی دوجمله‌ای با میانگین μ=n.p و واریانس σ^2=n.p.q باشد، هنگامی‌که n به‌سمت بینهایت میل‌کند، توزیع متغیر تصادفی $z=\frac{x-np}{\sqrt{npq}}$ به توزیع نرمال استاندارد میل می‌کند.

این تقریب، هنگامی‌که n بزرگ و p به 0.5 نزدیک باشد، بسیار خوب است.

تقریب توزیع دوجمله‌ای با توزیع پواسون و نرمال

برای تشخیص اینکه چه موقع می‌توان از توزیع نرمال یا توزیع پواسون برای تقریب توزیع دوجمله‌ای استفاده کرد، به‌عنوان یک قاعده سرانگشتی، ابتدا مقادیر np و nq را محاسبه کرده و سپس اگر:

- ✓ n بزرگ و p به 0.5 نزدیک باشد و در نتیجه np و nq، هر دو بزرگتر از 5 باشند، آنگاه می‌توان با تقریب خوبی توزیع دوجمله‌ای را با **توزیع نرمال** تقریب زد.
- ✓ n بزرگ و p به صفر نزدیک باشد؛ بطوریکه که رابطه $np \leq 10$ برقرار باشد؛ آنگاه می‌توان با تقریب خوبی توزیع دوجمله‌ای را با **توزیع پواسون** تقریب زد.

((به‌طورکلی، وقتی $n \geq ۲۰$ و $p < ۰.۰۵$ باشد، توزیع پواسون، تقریب خوبی برای احتمال‌های دوجمله‌ای می‌باشد.))

توزیع لگ‌نرمال

- هنگامی که تبدیل لگاریتم طبیعی داده‌ها توزیع نرمال داشته باشد، داده‌های اصلی دارای توزیع *lognormal* خواهند بود.

اگر
$$X\sim N(\mu,\sigma^2)$$
$$Y=e^X \rightarrow \ln(Y)=X$$

$$\leftarrow Y\sim LN(\mu,\sigma^2)$$

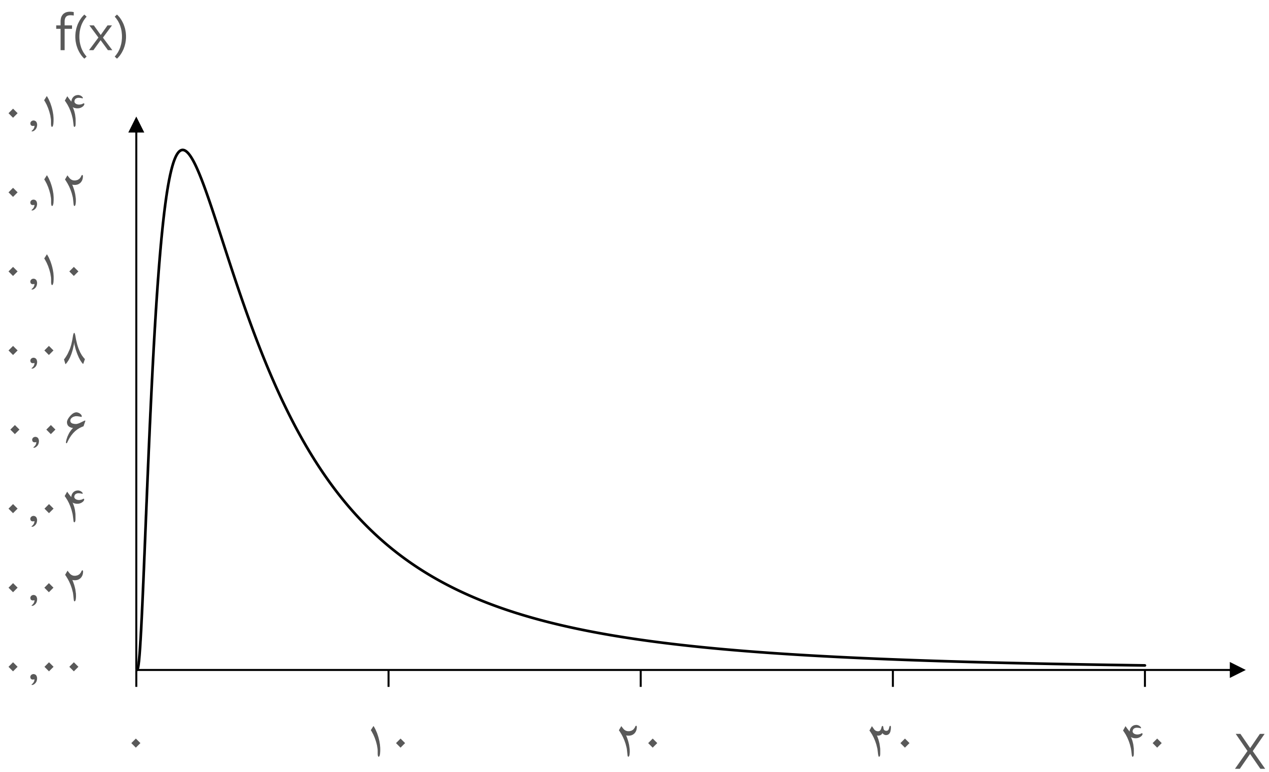
- تابع چگالی احتمال لگ‌نرمال برای y نامنفی به صورت زیر است؛

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

توجه: پارامترهای توزیع لگ‌نرمال، میانگین و واریانس این توزیع نیستند؛ بلکه تحت عنوان پارامترهای شکل و مقیاس شناخته می‌شوند.

توزیع لگ‌نرمال

- توزیع احتمالِ یک توزیع لگ‌نرمال با پارامترهای شکل و مقیاس به ترتیب برابر با ۵ و ۱



توزیع لگ‌نرمال

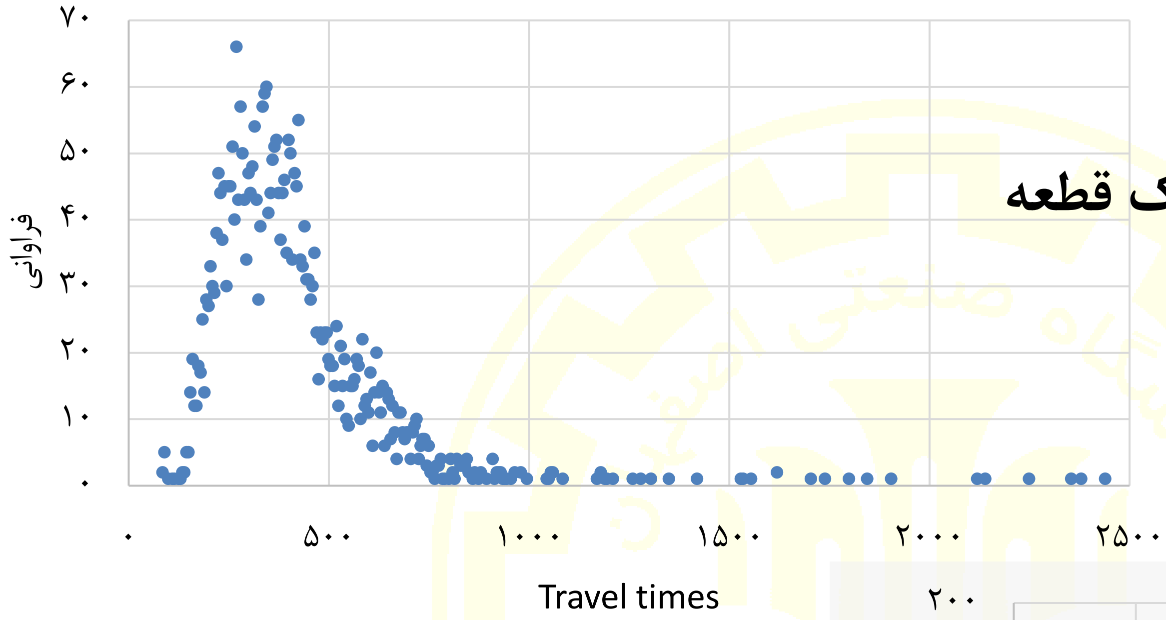
- یکی از مزایای این توزیع، نامنفی بودن و دنباله کشیده (چوله به راست) آن است که باعث می‌شود در مدلسازی پدیده‌های فیزیکی نسبت به توزیع نرمال ارجحیت داشته باشد.
- زیرا توزیع نرمال مقادیرمنفی نیز اختیار می‌کند و همچنین نقاط پرت کمی دارد.
- برای مثال، از توزیع لگ‌نرمال در مطالعات حمل‌ونقلی مربوط به بررسی میزان قابل رؤیت بودن تابلوها و علائم روی سطح روسازی (مثلا در زمان بارندگی) استفاده شده‌است.
- متغیر تصادفی: دورترین فاصله‌ای که علائم روی روسازی توسط راننده قابل مشاهده باشد.

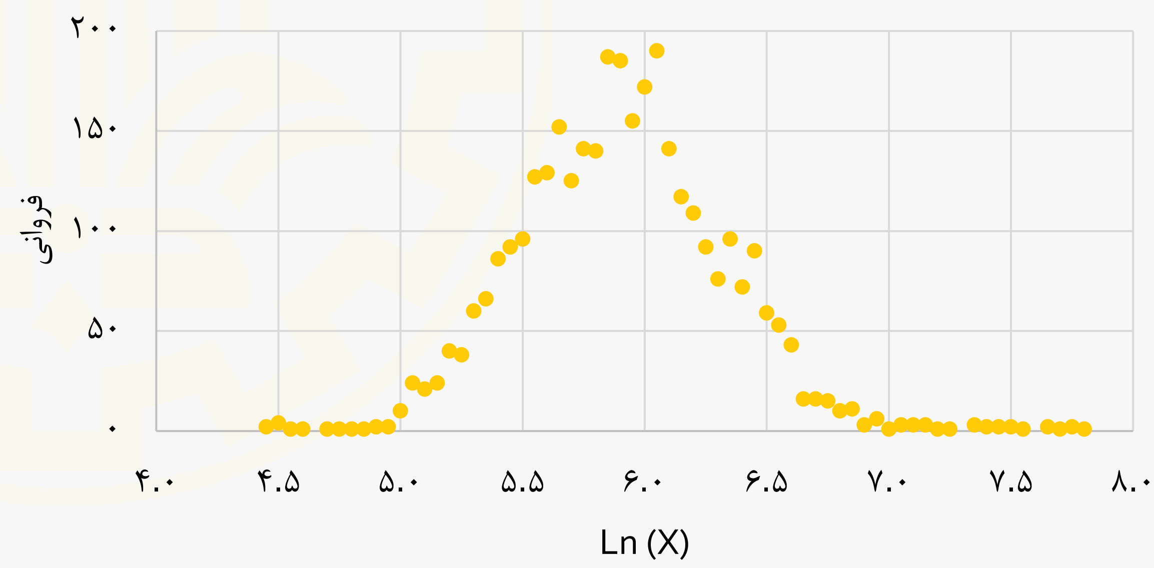
توابع چگالی احتمال و توزیع های پیوسته

مدرس: محمدمهدی بشارتی

توزیع لگ‌نرمال

○ داده‌های زمان سفر اتوبوس‌ها در یک قطعه





توزیع نمایی (Exp)

اگر فراوانی وقوع یک رخداد، فرض‌های زیر را ارضا نماید، آنگاه توزیع نمایی مدل مناسبی برای نمایش بازه‌های زمانی میان وقوع رخدادهای متوالی خواهد بود.

1. بتوان فراوانی وقوع پدیده موردنظر در یک بازه زمانی معین را به صورت فرآیند پواسون با پارامتر λ پیروی می‌کند، مدل کرد

$$(f(N(t) = X) = P(X = x) = (\frac{e^{\lambda t}\,\lambda t^{x}}{x!}))$$

1. میانگین مدل پواسون در طول زمان تغییر نکرده و وقایع با نرخ ثابت λ رخ دهند.
2. فراوانی وقوع پدیده موردنظر در یک بازه زمانی بسیارکوتاه، حداکثر برابر ۱ باشد (امکان وقوع همزمان وجود نداشته باشد).
3. وقوع پدیده موردنظر در دوره‌های زمانی مختلف، مستقل از هم باشد.

توزیع نمایی (Exp)

به‌عبارت دیگر، در یک فرآیند پواسون، اگر متغیر X نمایانگر تعداد رخداد در یک بازه زمانی معین باشد و متغیر t را به‌صورت زمان لازم برای مشاهده اولین رخداد تعریف کنیم؛

فرض کنید λ نمایانگر میانگین تعداد رخداد در واحد زمان باشد.

پس توزیع پواسون متناظر دارای میانگین $\mu = \lambda t$ در در هر بازه t خواهد بود.

$$P(X = x) = (\frac{e^{\lambda t}\ \lambda t^{x}}{x!})$$

احتمال اینکه طی زمان ۰ تا t هیچ رخدادی رخ ندهد (x=0) برابر است با؛

$$P(X = 0) = e^{-\lambda t}$$

اکنون متغیر تصادفی را به صورت زیر تعریف می‌کنیم

T: مدت زمان سپری شده تا مشاهده اولین رخداد

احتمال مشاهده رخداد بعد از زمان t، برابر است با احتمال عدم وقوع رخداد تا زمان t (یعنی از ۰ تا t)

$$P(T > t) = P(X = 0) = e^{-\lambda t}$$

توزیع نمایی (Exp)

احتمال مشاهده رخداد بعد از زمان t، برابر است با احتمال عدم وقوع رخداد تا زمان t (یعنی از ۰ تا t)

$$P(T>t)=P(X=0)=e^{-\lambda t} \xrightarrow[(T\leq t)]{\text{مکمل } (T>t)} P(T\leq t)=1-P(T>t)=1-e^{-\lambda t}$$

$(T\leq t)$ همان تعریف تابع توزیع است که می‌توان آن را به صورت زیر نوشت؛

$$F(t)=\begin{cases}0 & t<0\\ 1-e^{-\lambda t} & t\geq 0\end{cases}$$

با مشتق‌گیری از این تابع توزیع، می‌توان تابع چگالی را نیز به دست آورد؛

$$f(t)=\lambda e^{-\lambda t} \quad ,t>0$$

این تابع چگالی معروف به **توزیع نمایی** است (به جای متغیر T، هر متغیر دیگری مانند X را می‌توان به کار برد).

توزیع نمایی (Exp)

بدین‌ترتیب، **تابع چگالی احتمال توزیع نمایی** برای متغیر $t \geq 0$ به شکل روبروست

$$f(t) = \lambda e^{-\lambda t}$$

λ : میانگین تعداد وقوع رخداد موردنظر در واحد زمان

$$f(t) = \frac{1}{\beta} e^{-\frac{t}{\beta}}$$

$\beta = \frac{1}{\lambda}$: میانگین زمان انتظار میان دو رخداد متوالی

t: متغیر تصادفی مدت زمان سپری شده

توزیع نمایی معمولاً به صورت $EXP(\lambda)$ یا $EXP(\frac{1}{\lambda})$، یا $EXP(\beta)$ نمایش داده می‌شود.

توزیع نمایی (Exp)

- منحنی توزیع نمایی دارای دنباله کشیده بوده و معمولاً نسبت به داده‌های دارای توزیع نرمال، نقاط پرت بیشتری دارد.
- از توزیع نمایی در حوزه حمل‌ونقل، اغلب برای **مدلسازی زمان رسیدن وسایل‌نقلیه** و **عابرین در تقاطعات** و سایر نقاط مورد نظر استفاده می‌شود.

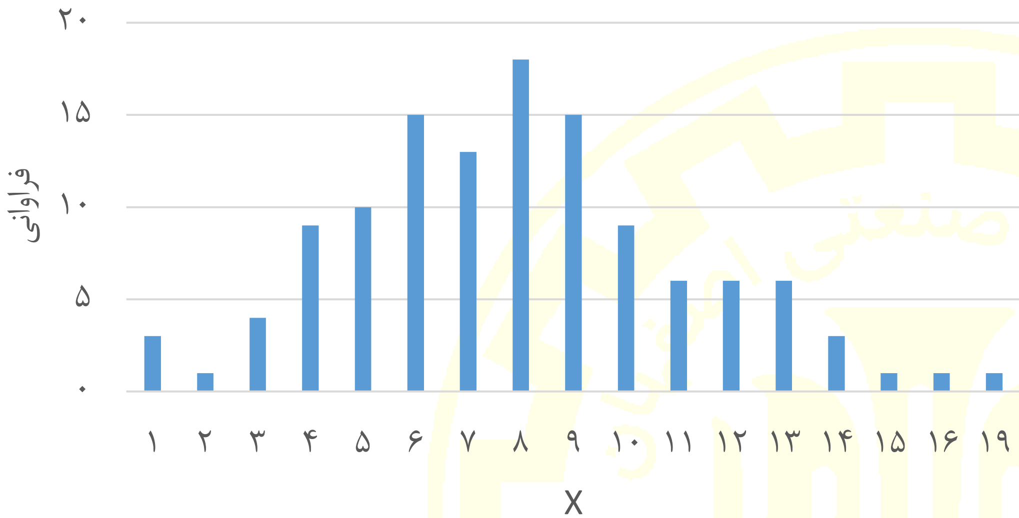
❖ علاوه بر محاسبه زمان مشاهده اولین پیروزی، یا زمان میان دو پیروزی متوالی، از توزیع نمایی در **تئوری صف** و **تئوری قابلیت اطمینان** نیز به‌طور گسترده‌ای استفاده می‌شود. برای مثال،

❖ اگر در یک سیستم صف، ورود مشتری‌ها طبق فرآیند پواسون باشد، می‌توان گفت زمان میان دو ورود متوالی از توزیع نمایی پیروی می‌کند و برعکس.

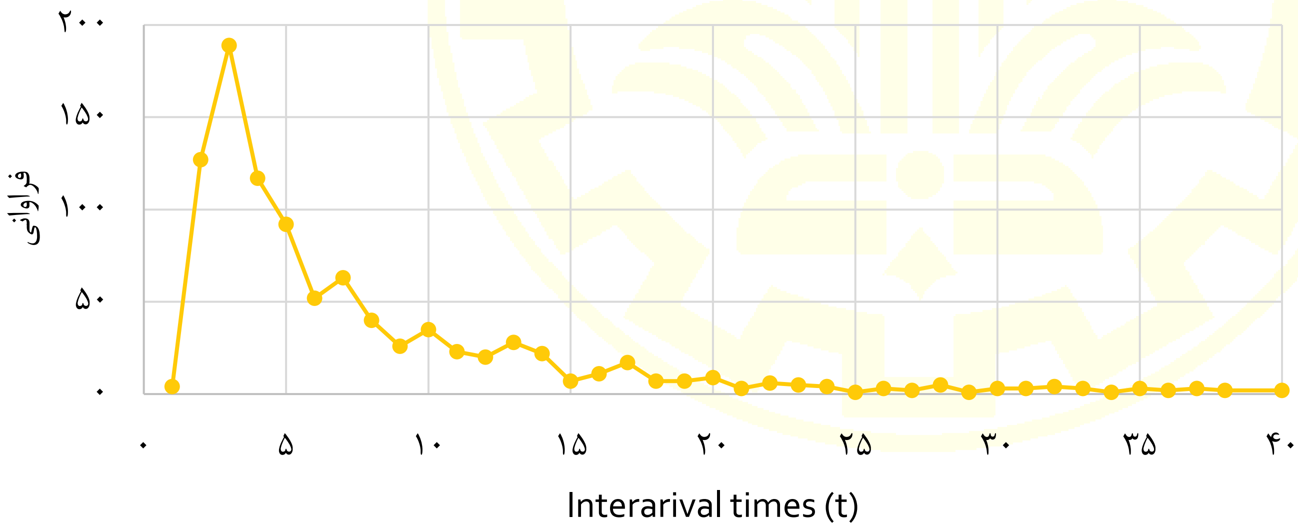
❖ این نوع اطلاعات برای طراحی ایستگاه‌های حمل و نقل همگانی و مکانیابی ایستگاه‌ها ضروری هستند.

توابع چگالی احتمال و توزیع های پیوسته

آماربرداری خودروهای ورودی به یک بزرگراه از یک رمپ



تعداد خودروهای ورودی در بازه های ۱ دقیقه‌ای



مدت زمان سپری شده میان ورود هر دو خودرو متوالی

توزیع نمایی (Exp)

مثال:

تجربه نشان داده‌است که تعداد اتومبیل‌هایی که وارد یک پارکینگ می‌شوند، دارای توزیع پواسون با نرخ 20 دستگاه در یک ساعت است($\lambda=20$). اگر پارکینگ ساعت 6 صبح باز شود، مطلوبست محاسبه احتمال اینکه:

الف) حداقل 5 دقیقه طول بکشد تا اولین اتومبیل به پارکینگ مراجعه کند.

ب) تا ساعت $6:10$ صبح، اولین اتومبیل وارد پارکینگ شود.

ج) اگر تا ساعت 7 صبح اتومبیلی وارد پارکینگ نشده باشد، احتمال اینکه تا ساعت $7:05$ اتومبیلی وارد پارکینگ شود چقدر است؟

توزیع نمایی (Exp)

مثال:

ابتدا تعداد موفقیت در دقیقه را حساب می‌کنیم: $\frac{20}{60}=\frac{1}{3}$ ، اگر متغیر t نشان‌دهنده مدت زمان انتظار برای اولین ورود اتومبیل به پارکینگ باشد، t دارای توزیع نمایی با پارامتر $\frac{1}{3}$ خواهد بود.

$$f(t)=P(T=t)=\frac{1}{3}e^{-\frac{1}{3}t}$$

الف) احتمال اینکه حداقل 5 دقیقه طول بکشد تا اولین اتومبیل به پارکینگ مراجعه کند.

$$f(T>5)=\int_5^\infty \frac{1}{3}e^{-\frac{1}{3}t}dt=e^{-\frac{5}{3}}$$

ب) احتمال اینکه تا ساعت *6:10* صبح، اولین اتومبیل وارد پارکینگ شود.

$$f(T<10)=\int_0^{10}\frac{1}{3}e^{-\frac{1}{3}t}dt=\frac{1}{3}(1-e^{-\frac{10}{3}})$$

توزیع نمایی (Exp)

مثال:

ج) یکی از مهم‌ترین ویژگی‌های توزیع نمایی، **بدون حافظه بودن این توزیع** است. بدین معنا که در این توزیع، می‌توان بدون درنظر گرفتن اتفاقات گذشته، هر زمانی را بعنوان مبدا محاسبات درنظر گرفت.

فرض کنید زمان وقوع یک رویداد، متغیر تصادفی با توزیع نمایی باشد اگر این رویداد تا لحظه معینی، مثلا t_1 رخ نداده باشد، می‌توان از این زمان صرف‌نظر کرد و مبدا زمان را به این لحظه (لحظه t_1 بجای صفر) انتقال داد. بدین‌ترتیب، احتمال اینکه تا زمان t_1+t_2 رخداد موردنظر رخ بدهد، برابر است با احتمال اینکه تا زمان t_2 رخداد موردنظر رخ‌دهد:

$$P(T > t_1 + t_2 \,|\, T > t_1) = P(T > t_2)$$

براین اساس، می‌توان گفت در حالیکه تا ساعت 7 هیچ اتومبیلی وارد پارکینگ نشده باشد، احتمال اینکه تا ساعت $7{:}05$ اتومبیلی وارد پارکینگ شود، با احتمال اینکه در همان 5 دقیقه اول آغاز به کار، یک اتومبیل وارد پارکینگ شود برابر است.

توزیع گاما

فرآیند ساخت تابع توزیع گاما مشابه توزیع نمایی است. با این تفاوت که در توزیع گاما به دنبال "زمان سپری‌شده تا مشاهده α اُمین رویداد" هستیم.

فرض کنید متغیر تصادفی X را به صورت زیر تعریف کرده‌ایم؛

X : زمان سپری شده بین α مرتبه رخداد متوالی (**زمان انتظار تا مشاهده α اُمین رویداد**)

متغیر تصادفی X دارای توزیع گاما با پارامترهای α و β بوده و تابع چگالی احتمال آن برای $X \geq 0$ به صورت زیر است (زمان سپری‌شده برای مشاهده α اُمین رویداد) :

$$f(x;\ \alpha,\beta) = \frac{\beta^{\alpha}}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$$

α، تعداد رویدادهای مشاهده شده

به جای β می‌توان از λ استفاده کرد (فراوانی وقوع رخداد موردنظر در واحد زمان).

بدین ترتیب، **فرم دیگری** از تابع چگالی توزیع گاما (**Gamma Distribution – Rate Parameterization**)؛

$$f(x;\ \alpha,\lambda) = \frac{\lambda e^{-\lambda x} (\lambda x)^{\alpha-1}}{\Gamma(\alpha)}$$

توزیع گاما

فرم دیگری از تابع چگالی توزیع گاما به صورت زیر تعریف می‌گردد (**Gamma Distribution – Scale Parameterization**)؛

زمان سپری شده بین α مرتبه رخداد متوالی (زمان سپری‌شده تا مشاهده α اُمین رویداد)

متغیر تصادفی X دارای توزیع گاما با پارامترهای α و β بوده و تابع چگالی احتمال آن برای $X \geq 0$ به صورت زیر است (زمان سپری‌شده برای مشاهده α اُمین رویداد) :

$$f(x;k,\theta) = \frac{x^{\alpha-1}e^{\frac{-x}{\theta}}}{\Gamma(\alpha)\,\theta^{\alpha}}$$

α، تعداد رویدادهای مشاهده شده

θ، معکوس λ است. یعنی متوسط مدت زمانی‌است که باید سپری شود تا رخداد بعدی (مثلا رسیدن خودرو بعدی به تقاطع) مشاهده شود (زمان مورد انتظار بین دو رخداد متوالی).

توجه: α، پارامتر شکل (Shape parameter)

λ، پارامتر نرخ (Rate parameter)

θ، پارامتر مقیاس (Scale parameter)

توزیع گاما

- ✓ توزیع‌های **پواسون**، **نمایی**، و **گاما (اِرلانگ)** همگی به فرآیند پواسون از منظرهای مختلف نگاه می‌کنند؛
  - ○ **پواسون:** تعداد رویدادها در یک بازه زمانی معین
  - ○ **نمایی:** زمان مشاهده اولین رویداد ($\alpha = 1$)
  - ○ **گاما:** زمان مشاهده α-امین رویداد

- ✓ مهندسین حمل‌و نقل از توزیع گاما برای مدلسازی **سرفاصله زمانی** یا **مکانی (بر حسب زمان) بین وسایل‌نقلیه** استفاده می‌کنند.

- برای یک متغیر تصادفی X که از توزیع گاما(α , λ) پیروی می‌کند:

$$E(X) = \frac{\alpha}{\lambda}$$

$$Var(X) = \frac{\alpha}{\lambda^2}$$

توزیع اِرلانگ

- ✓ در توزیع گاما، α می‌تواند یک **عدد حقیقی مثبت** باشد.
- ✓ هنگامی‌که α یک **عدد صحیح** است، مقدار Γ(α) برابر با !(α−۱)خواهد بود.
- ✓ در این شرایط، به توزیع گاما، توزیع اِرلانگ ((X ~ Er(α , λ) می‌گویند (یعنی شرایطی که α یک عدد صحیح است).
- ✓ تابع توزیع تجمعی توزیع ارلانگ با پارامترهای α و λ (مکمل CDFپواسون است؟):

$$F_X(x;\alpha,\lambda) = P(X \leq x) = 1 - \sum_{i=0}^{\alpha-1} \frac{e^{-\lambda x}(\lambda x)^i}{i!}$$

این تابع، احتمال این‌که «رخداد α-ام، قبل از زمان x رخ دهد» را اندازه‌گیری می‌کند (X: **زمان انتظار برای α رخداد**).

تابع چگالی احتمال این توزیع به صورت‌های زیر قابل تعریف است؛

$$f(x;k,\theta) = \frac{x^{\alpha-1}e^{\frac{-x}{\theta}}}{(\alpha-1)!\ \theta^{\alpha}}$$

$$f(x;\ \alpha,\lambda) = \frac{\lambda^{\alpha}}{(\alpha-1)!}x^{\alpha-1}e^{-\lambda x}$$

$$f(x;\ \alpha,\beta) = \frac{\beta^{\alpha}}{(\alpha-1)!}x^{\alpha-1}e^{-\beta x}$$

توزیع اِرلانگ

به عبارت دیگر، اگر متغیرهای تصادفی T_i دارای توزیع‌های مستقل و مشابه نمایی با پارامتر λ باشند؛

$$T_i \sim EXP(\lambda)$$

آنگاه؛ متغیر Y (مجموع α متغیر نمایی مستقل با نرخ λ) از توزیع اِرلانگ با پارامترهای λ و α پیروی کند؛

$$Y = \sum_{i=1}^{\alpha} T_i \sim Erlang(\alpha, \lambda) = Gamma(\alpha, \lambda)$$

اگر فراوانی وقوع رخداد موردنظر در واحد زمان، از توزیع پواسون با پارامتر λ پیروی کند؛ آنگاه، مدت‌زمانی که طول می‌کشد تا n مورد از رخداد موردنظر اتفاق بیفتد (فاصله زمانی میان وقوع ۱+n رخداد متوالی) از توزیع ارلانگ با پارامتر λ و $\alpha=n$ پیروی خواهد کرد.

پس توزیع ارلانگ، توزیع مجموع n متغیر تصادفی مستقل و هم‌توزیع (هر یک دارای توزیع نمایی) است (برای حالتی که α به صورت یک عدد صحیح باشد).

خلاصه

نحوه ارتباط توزیع‌های پواسون، نمایی، گاما (ارلانگ)

اگر

- رویدادها براساس یک فرآیند پواسون با نرخ λ اتفاق بیفتند $X \sim Poisson(\lambda)$
- مدت زمان میان هر دو روخداد متوالی (زمان انتظار) از توزیع نمایی پیروی کند $T_i \sim EXP(\lambda)$

آنگاه؛

- مجموع n زمان انتظار از توزیع گاما با پارامترهای n و λ پیروی می‌کند $Y = \sum_{i=1}^{n} T_i = Gamma(n,\lambda)$

توزیع ارلانگ

مثال: فرض کنید تعداد خودروهایی که به یک پارکینگ عمومی مراجعه می‌کنند، دارای توزیع پواسون با میانگین ۳۰ خودرو در ساعت می‌باشد. مطلوب است احتمال این‌که از لحظه شروع به کار، مأمور پارکینگ حداقل ۵ دقیقه منتظر بماند تا دومین خودرو وارد پارکینگ شود؟

پاسخ.

در صورت مسئله، متوسط تعداد خودرو ورودی به پارکینگ در واحد ساعت بیان شده، اما متغیر تصادفی مسئله، زمان انتظار بر حسب دقیقه می‌باشد. بنابراین، متوسط تعداد خودروهایی که در دقیقه وارد پارکینگ می‌شوند برابر است با ($\frac{30}{60}=\frac{1}{2}$).

بدین‌ترتیب، متغیر تصادفی X دارای توزیع گاما (ارلانگ) با پارامترهای α=۲ و $\lambda=\frac{1}{2}$ می‌باشد. بنابراین احتمال این‌که مأمور پارکینگ حداقل ۵ دقیقه منتظر بماند تا دومین اتومبیل وارد پارکینگ شود برابر است با

$$P(X>5)=1-F_X(x;\ \alpha,\lambda)=\sum_{i=0}^{\alpha-1}\frac{e^{-\lambda x}(\lambda x)^i}{i!}=\frac{e^{-\frac{۵}{۲}}\times(\frac{۵}{۲})^{۰}}{۰!}+\frac{e^{-\frac{۵}{۲}}\times(\frac{۵}{۲})^{۱}}{۱!}=\frac{۷}{۲}e^{-\frac{۵}{۲}}$$

توزیع بتا (β)

اگر متغیر X از توزیع بتا با پارامترهای $(\alpha$ و $\beta)$ پیروی کند؛ تابع چگالی احتمال آن به صورت زیر است.

$$f(x;\alpha,\beta)=\frac{\Gamma(\beta+\alpha)}{\Gamma(\alpha)\Gamma(\beta)}x^{\alpha-1}(1-x)^{\beta-1} \qquad 0<x<1$$

- مقادیر مختلف α و β اشکال مختلفی از توابع چگالی احتمال برای توزیع بتا را تولید می‌کنند.
- اگر α و β **اعدا طبیعی** باشند، $\Gamma(\alpha)$ و $\Gamma(\beta)$ را می‌توان به صورت فاکتوریل نوشت و تابع بالا را به شکل ساده‌تری نوشت.
- مثلا اگر، $\alpha=m+1$ و $\beta=n-m+1$ در نظر بگیریم؛ توزیع بتا به صورت زیر در می‌آید.

$$f(x)=(n+1)\frac{n!}{m!\,(n-m)!}x^{m}(1-x)^{n-m}$$

- این همان توزیع بتای مزدوج دوجمله‌ای است.
- در واقع، توزیع بتا به عنوان یک مدل برای **احتمال موفقیت (p) در یک توزیع دوجمله‌ای** استفاده می‌گردد.
- یعنی در اینجا، متغیر تصادفی X، مشابه با پارامتر p در توزیع دوجمله‌ای است.

توزیع بتا (β)

- بتا توزیع پیشین مزدوج (conjugate prior) برای پارامتر p در توزیع دوجمله‌ای است.
- از توزیع بتا می‌توان به عنوان یک توزیع پیشینِ توزیع دو جمله‌ای برای برآورد بیزی پارامتر احتمال موفقیت استفاده کرد.

  - اگر $X \sim Binomial\ (n,p)$
  - آنگاه توزیع پیشین برای p عبارتست از: $p \sim Beta\ (\alpha,\beta)$
  - و بدین ترتیب، توزیع پسین نیز بتا خواهد بود.

تابع گاما (برای اعداد صحیح)

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

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$$

$$\Gamma(1) = 1$$

$$\Gamma(2) = 1$$

تابع بتا برحسب گاما

$$B\ (x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$$

توزیع کای‌دو (Chi-2)

- اگر $X_1, X_2, \ldots, X_n$ متغیرهای تصادفی مستقل نرمال استاندارد ((N(0,1)) باشند؛
- آنگاه متغیر تصادفی $X = \sum_{i=1}^{n} X_i^2$ دارای توزیع گاما با $\alpha = \frac{n}{2}$ و $\theta = 2$ خواهد بود ($X \sim Gamma\ (\alpha = \frac{n}{2}\ , \theta = 2)$).
- همچنین، این متغیر تصادفی از توزیع کای‌دو با n درجه آزادی پیروی می‌کند ($X \sim \chi^2\ (df = n)$).
- بنابراین، این توزیع، **توزیع کای‌دو با n درجه‌آزادی** نامیده می‌شود.
- **توزیع کای‌دو** حالت خاصی از **توزیع گاما** است ($X \sim\ Gamma\ (\alpha = \frac{n}{2}\ , \theta = 2)$).
- تابع چگالی احتمال مربوط به توزیع کای‌دو با df درجه‌آزادی به صورت زیر است:

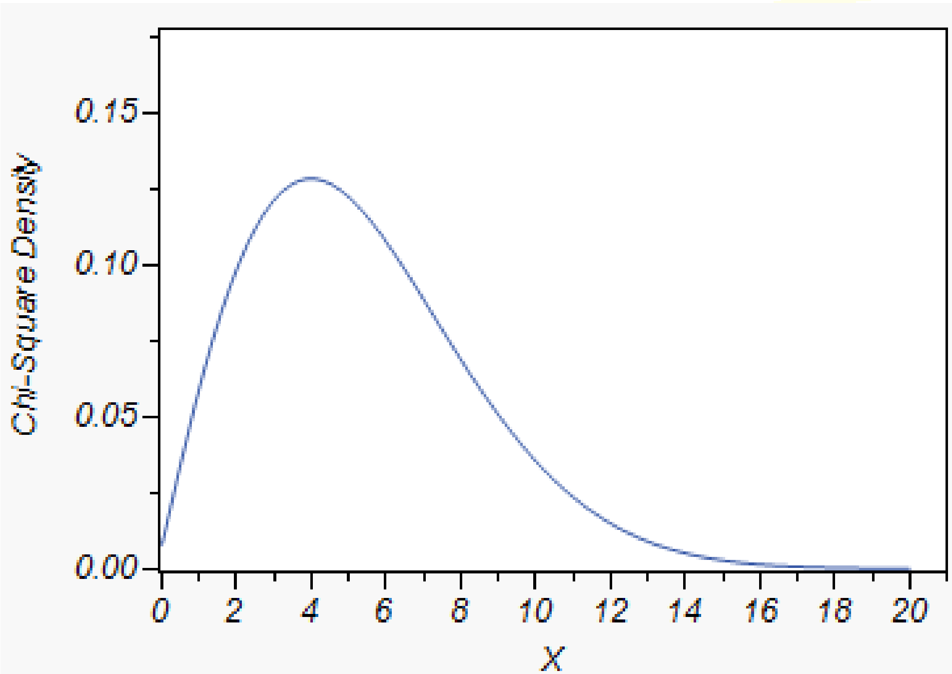
$$f(x) = \frac{(x)^{\frac{df}{2}-1}(e)^{\frac{-x}{2}}}{\Gamma\left(\frac{df}{2}\right)\ (2)^{\frac{df}{2}}} \quad , \qquad x > 0$$

$$E(X) = n$$

$$Var(X) = 2n$$

توزیع کای‌دو (Chi-2)

توزیع chi-square با ۵ درجه آزادی



- توزیع کای‌دو عمدتا برای آزمون آماری مورد استفاده قرار می‌گیرد.
- سایر کاربردها؛
  - در مدل‌های بیزی نیز استفاده می‌شود.
  - در مدل‌های خطی تعمیم‌یافته هم ظاهر می‌شود.
  - در نظریه احتمال و فرآیندهای تصادفی نیز کاربرد دارد.

- توزیع کای‌دو «چوله به راست» است. اما؛
- با افزایش درجات آزادی آن (n)، متقارن‌تر شده و به توزیع نرمال میل می‌کند.
- بنابراین می‌توان آن را با توزیع نرمال با میانگین n و واریانس 2n تقریب زد.
- به عبارت دیگر، متغیر تعریف شده به صورت روبرو از توزیع نرمال استاندارد پیروی می‌کند؛

$$\frac{X-n}{\sqrt{2n}} \to N(0,1)$$

توزیع کای‌دو (Chi-2)

- توزیع کای‌دو رایج‌ترین توزیع برای مدل‌سازی **توزیع واریانس برآورد شده از داده‌های دارای توزیع نرمال** است.

- به عبارت دیگر، اگر متغیرهای (X_i) مستقل دارای توزیع نرمال با میانگین μ و واریانس σ^2 باشند، نسبت زیر دارای توزیع کای‌دو با $n-1$ درجه‌آزادی است.

$$\frac{(n-1)S^2}{\sigma^2} = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{\sigma^2} \quad \sim \; \chi^2 \; (df = n-1)$$

- در فصل های آینده خواهیم دید که چگونه با استفاده از این توزیع، درمورد واریانس جامعه استنباط می‌شود.

- سایر کاربردهای این نتیجه: آزمون‌های نیکویی برازش، و آزمون‌های استقلال.

توزیع t

- اگر Z و Y دو متغیر تصادفی مستقل بوده و
- Z دارای توزیع نرمال استاندارد و
- Y دارای توزیع کای‌دو با df درجه آزادی باشند؛
- آنگاه متغیر تصادفی T که به‌صورت زیر تعریف می‌شود؛ دارای توزیع t با df درجه آزادی خواهد بود؛

$$T = \frac{\mathrm{Z}}{\sqrt{\mathrm{Y}/\mathrm{df}}} \quad \sim \quad t_{df}$$

تابع چگالی احتمالِ این متغیر تصادفی به‌صورت زیر است (دامنه روی کل اعداد حقیقی است)؛

$$\mathrm{f_{df}(x)} = \frac{\Gamma\left(\frac{(\mathrm{df}+۱)}{۲}\right)}{\Gamma\left(\frac{\mathrm{df}}{۲}\right)\sqrt{\pi.\mathrm{df}}}\left(۱+\frac{\mathrm{x}^{۲}}{\mathrm{df}}\right)^{-(\mathrm{df}+۱)/۲}$$

$$E(X) = 0$$

$$Var(X) = \frac{df}{df-2}$$

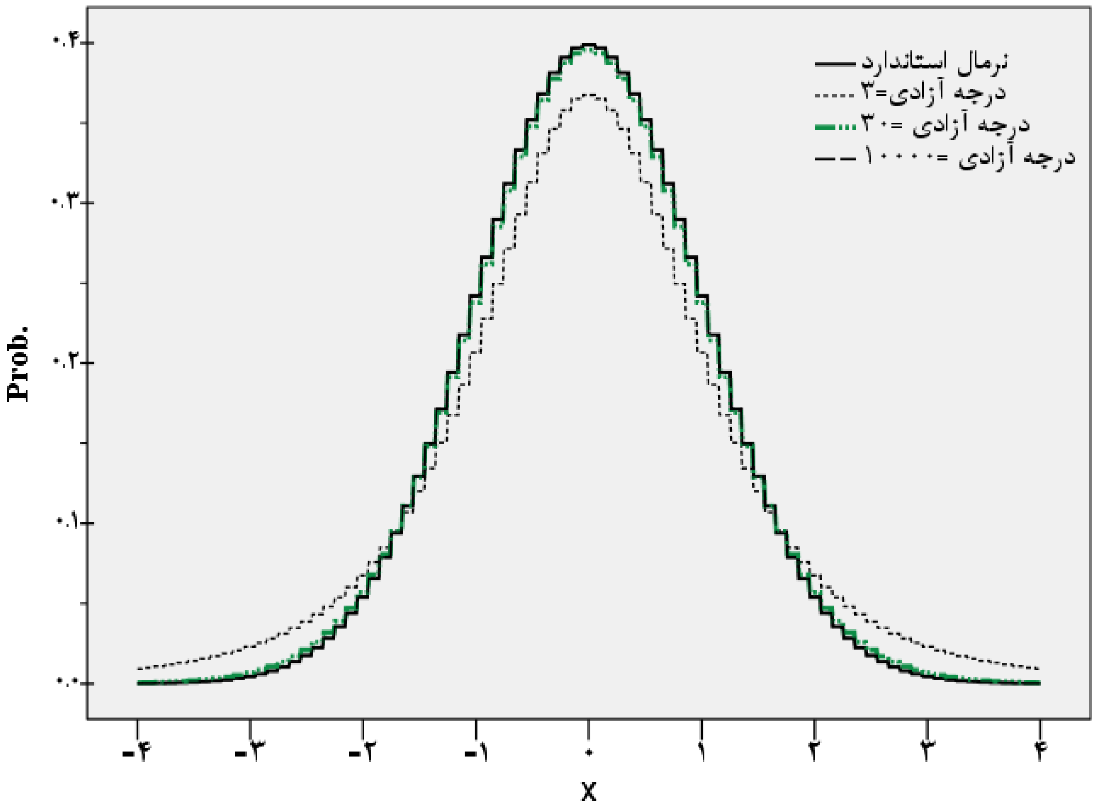
توزیع t

- توزیع t همانند توزیع نرمال استاندارد نسبت به $\mu=0$ متقارن بوده و به شکل زنگوله است.
- دنباله‌های این توزیع کشیده‌تر از دنباله‌های توزیع نرمال استاندارد می‌باشد.
- این توزیع برای انجام آزمون‌های آماری مورد استفاده قرار می‌گیرد.
- در شرایطی که واریانس جامعه نامعلوم باشد یا داده‌ها دارای دنباله‌های سنگین باشند (وجود داده‌های پرت)، استفاده از مدل مبتنی بر توزیع t مناسب‌تر از نرمال است.

توزیع t

- نمودارهای توزیع t را می‌توان به‌ازای درجات آزادی مختلف به دست آورد.
- با افزایش درجات آزادی، توزیع t به توزیع نرمال استاندارد نزدیک می‌شود.
- توزیع t با $df = \infty$ درجه‌آزادی بر توزیع نرمال استاندارد منطبق می‌شود.
- برای درجات آزادی بزرگ (مثلاً $df > 30$)، تفاوت آن با نرمال ناچیز می‌شود.

توزیع z و
توزیع t (به‌ازای ۳ ، ۳۰ و ۱۰۰۰۰ درجه‌آزادی)



توزیع t

- اگر $\bar{X}$و $S^۲$ به ترتیب، میانگین و واریانس یک نمونه تصادفی nتایی از یک جامعه نرمال با میانگین μ و واریانس $\sigma^۲$ باشند؛ آنگاه متغیر تصادفی T که به‌صورت زیر تعریف می‌شود؛ دارای توزیع t با df=n- ۱ درجه آزادی خواهد بود:

$$T=\frac{(\bar{X}-\mu)}{(S/\sqrt{n})} \sim t_{n-۱} \qquad S^2=\frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$$

مقادیر احتمال متناظر با برخی از صدک‌های پرکاربرد توزیع t، به ازای مقادیر مختلف df و α در قالب جداول آماده در کتب آماری ارائه می‌شود.

در این جدول‌ها، $t_{\alpha,df}$ مقداری از متغیر تصادفی t است که احتمال مشاهده مقداری بیشتر از آن در یک توزیع t با df درجه آزادی، برابر با α است.

✓ موارد کاربرد توزیع t: آزمون آماری میانگین یک یا دو جامعه، آزمون معناداری ضرایب رگرسیون.

توزیع فیشر (F)

- توزیع فیشر در بسیاری از آزمون‌های فرضیه آماری از جمله تحلیل‌واریانس، رگرسیون و آزمون برابری واریانس مورد استفاده قرار می‌گیرد.

- توزیع فیشر از توزیع نسبت **دو متغیر تصادفی مستقل کای‌دو** (که هرکدام بر تعداد درجه‌های آزادی خود تقسیم شده‌است) حاصل می‌شود.

- اگر X و Y دو متغیر تصادفی مستقل کای‌دو، به ترتیب با df_1 و df_2 درجه‌آزادی باشند، آماره زیر دارای توزیع فیشر با df_1 و df_2 درجه‌آزادی است.

$$X \sim \chi^2\,(df_1) \qquad Y \sim \chi^2(df_2)$$

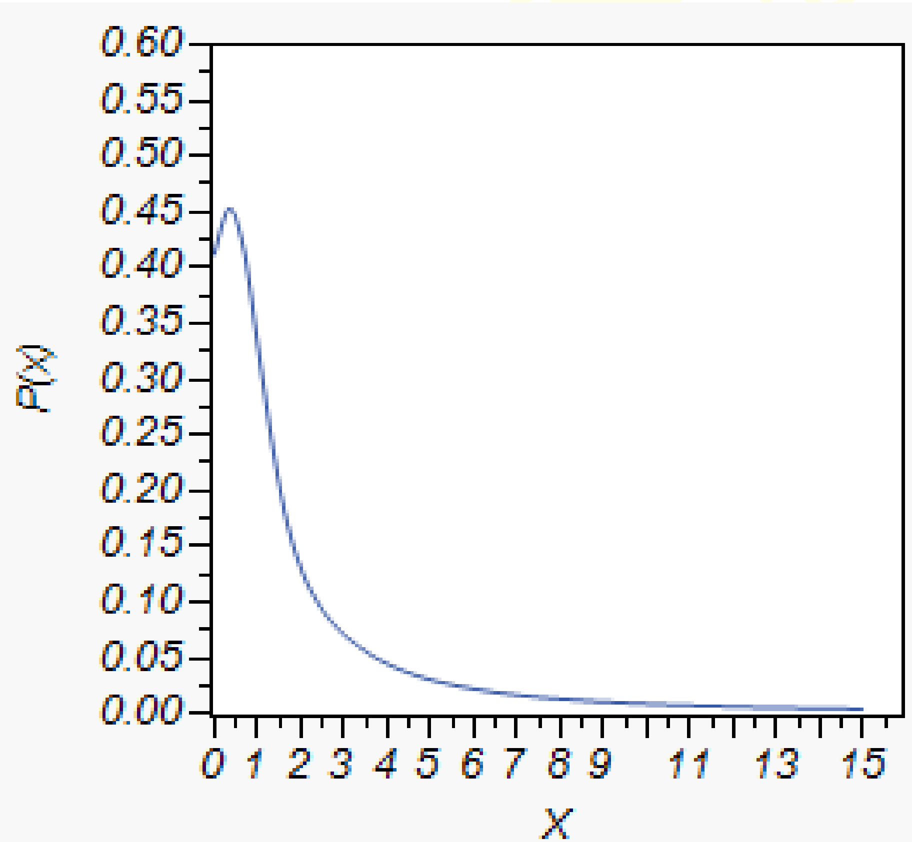
$$F = \frac{\left(\frac{X}{df_1}\right)}{\left(\frac{Y}{df_2}\right)} \sim F\,(df_1, df_2)$$

- **تفسیر:** توزیع فیشر نسبت ۲ برآورد مستقل و مقیاس شده از واریانس را ارائه می‌کند.
- برای مقایسه واریانس‌ها استفاده می‌شود (ANOVA)

توزیع فیشر (F)

این توزیع دارای ویژگی‌های زیر است:

- شکل توزیع فیشر، با دو عدد مثبت مشخص می‌گردد (درجه آزادی صورت و مخرج)
- نامتقارن و چوله به راست است.
- مقادیر متغیرتصادفی توزیع فیشر، نامنفی می‌باشد.

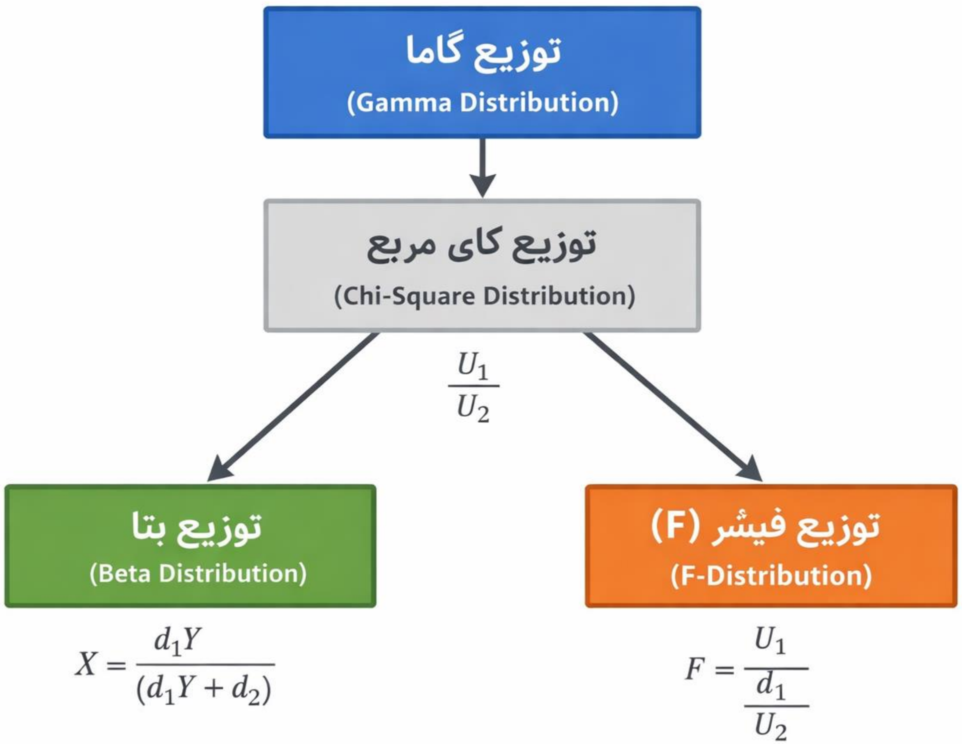


$df_1=5$, $df_2=2$

- **کاربردهای توزیع فیشر:**
- آزمون تساوی واریانس‌ها
- آزمون تحلیل واریانس‌ها (ANOVA)
- آزمون معناداری کلی مدل رگرسیون

توزیع فیشر (F)

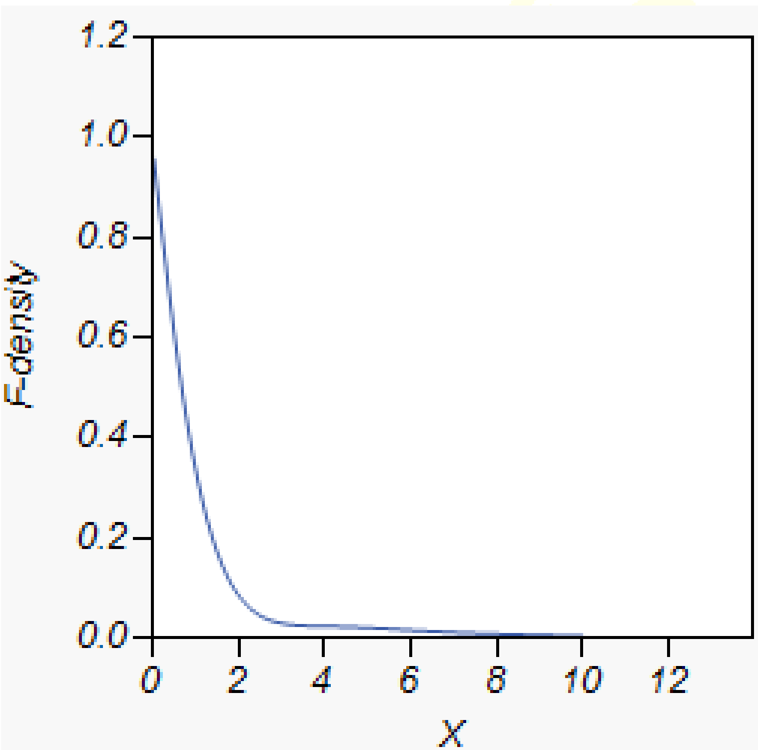
توزیع F از نسبت دو متغیر کای‌دو ساخته می‌شود که آن‌ها خود حالت خاصی از گاما هستند؛ بنابراین ریشه در گاما/بتا دارد.



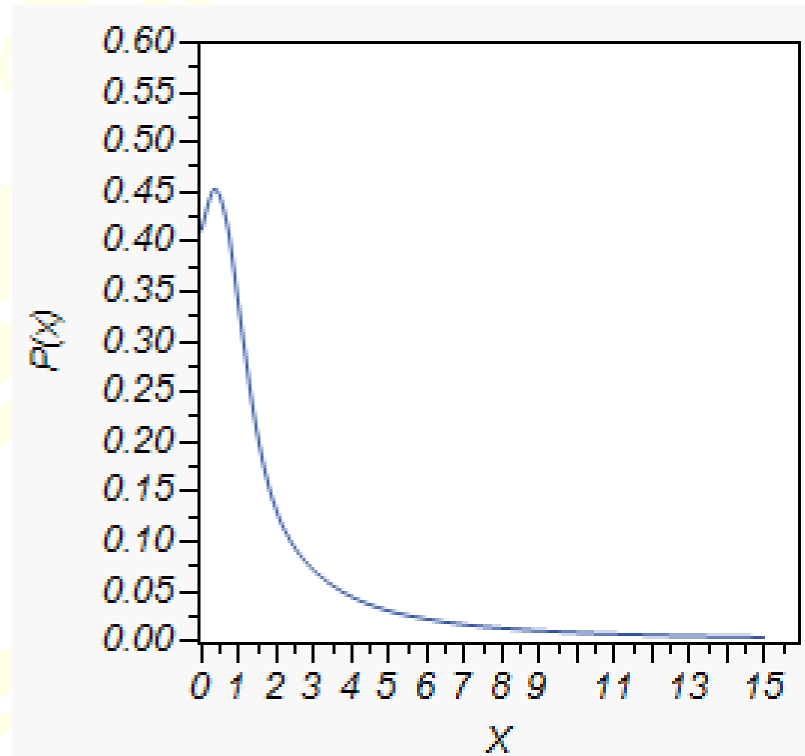
توابع چگالی احتمال و توزیع های پیوسته

توزیع فیشر (F)

توزیع فیشر غالبا به‌صورت $F(df_1 , df_2)$ نمایش داده می‌شود. توجه‌شود که ترتیب درجات آزادی برای توزیع فیشر، اهمیت داشته و نمی‌توان آن را جابجا نمود.



$df_1=2$, $df_2=5$



$df_1=5$, $df_2=2$

In general [edit]

- If $X \sim \chi^2_{d_1}$ and $Y \sim \chi^2_{d_2}$ (Chi squared distribution) are independent, then $\frac{X/d_1}{Y/d_2} \sim \mathrm{F}(d_1, d_2)$
- If $X_k \sim \Gamma(\alpha_k, \beta_k)$ (Gamma distribution) are independent, then $\frac{\alpha_2 \beta_1 X_1}{\alpha_1 \beta_2 X_2} \sim \mathrm{F}(2\alpha_1, 2\alpha_2)$
- If $X \sim \mathrm{Beta}(d_1/2, d_2/2)$ (Beta distribution) then $\frac{d_2 X}{d_1(1-X)} \sim \mathrm{F}(d_1, d_2)$
- Equivalently, if $X \sim F(d_1, d_2)$, then $\frac{d_1 X/d_2}{1 + d_1 X/d_2} \sim \mathrm{Beta}(d_1/2, d_2/2)$.
- If $X \sim F(d_1, d_2)$, then $\frac{d_1}{d_2} X$ has a beta prime distribution: $\frac{d_1}{d_2} X \sim \beta'\left(\frac{d_1}{2}, \frac{d_2}{2}\right)$.
- If $X \sim F(d_1, d_2)$ then $Y = \lim_{d_2 \to \infty} d_1 X$ has the chi-squared distribution $\chi^2_{d_1}$
- $F(d_1, d_2)$ is equivalent to the scaled Hotelling's T-squared distribution $\frac{d_2}{d_1(d_1 + d_2 - 1)} \mathrm{T}^2(d_1, d_1 + d_2 - 1)$.
- If $X \sim F(d_1, d_2)$ then $X^{-1} \sim F(d_2, d_1)$.
- If $X \sim t_{(n)}$ — Student's t-distribution — then:

$$X^2 \sim \mathrm{F}(1, n)$$
$$X^{-2} \sim \mathrm{F}(n, 1)$$

توزیع‌های دو متغیره

مواردی که تا اینجا بررسی شد، توزیع‌های یک‌متغیره بود.

در توزیع‌های دومتغیره به بررسی همزمان ۲ متغیر می‌پردازیم.

مثلا بررسی همزمان توزیع احتمال متغیر X و متغیر Y

توزیع‌های دو (چند) متغیره را توزیع مشترک متغیرهای تصادفی می‌نامند (پیوسته یا ناپیوسته).

توزیع‌های دو متغیره

نمایش توزیع مشترک دو متغیر تصادفی ناپیوسته (به صورت جدول)

X \ Y	y_1	y_2	…	y_m	جمع سطری ($p_{i.}$)
x_1	p_{11}	p_{12}	…	p_{1m}	$p_{1.}$
x_2	p_{21}	p_{22}	…	p_{2m}	$p_{2.}$
⋮	⋮	⋮	⋮	⋮	⋮
x_n	p_{n1}	p_{n1}	…	p_{nm}	$p_{n.}$
جمع ستونی ($p_{.j}$)	$p_{.1}$	$p_{.2}$	…	$p_{.m}$	۱

p_{ij}: احتمال اینکه به طور همزمان،

$X = x_i$ و $Y = y_j$

باشد، برابر است با؛

$$p_{ij} = P(X = x_i , Y = y_j)$$

$$p_{i.} = \sum_{j=1}^{m} p_{ij} \qquad i = 1,2, \dots , n$$

$$p_{.j} = \sum_{i=1}^{n} p_{ij} \qquad j = 1,2, \dots , n$$

توزیع‌های دو متغیره

نمایش توزیع مشترک دو متغیر تصادفی ناپیوسته (به صورت جدول)

Y / X	y_1	y_2	$\cdots$	y_m	توزیع حاشیه‌ای (X)
x_1	p_{11}	p_{12}	$\cdots$	p_{1m}	$p_{1.}$
x_2	p_{21}	p_{22}	$\cdots$	p_{2m}	$p_{2.}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
x_n	p_{n1}	p_{n1}	$\cdots$	p_{nm}	$p_{n.}$
جمع ستونی ($p_{.j}$)	$p_{.1}$	$p_{.2}$	$\cdots$	$p_{.m}$	۱

توزیع حاشیه ای X

$$P(X = x_i) = p_{i.}$$

توزیع تک متغیره X

توزیع‌های دو متغیره

نمایش توزیع مشترک دو متغیر تصادفی ناپیوسته (به صورت جدول)

Y \ X	y_1	y_2	…	y_m	جمع سطری $(p_{i.})$
x_1	p_{11}	p_{12}	…	p_{1m}	$p_{1.}$
x_2	p_{21}	p_{22}	…	p_{2m}	$p_{2.}$
⋮	⋮	⋮	⋮	⋮	⋮
x_n	p_{n1}	p_{n1}	…	p_{nm}	$p_{n.}$
توزیع حاشیه‌ای (Y)	$p_{.1}$	$p_{.2}$	…	$p_{.m}$	۱

توزیع حاشیه ای Y

توزیع تک متغیره Y

$$P(Y = y_j) = p_{.j}$$

توزیع‌های دو متغیره

احتمال‌های شرطی

$$P\left(X = x_j \middle| Y = y_j\right) = \frac{P(X = x_j, Y = y_j)}{P(Y = y_j)} = \frac{p_{ij}}{p_{.j}}$$

توزیع‌های احتمال شرطی X

X	$P\left(X \middle| Y = y_j\right)$
x_1	$\frac{p_{1j}}{p_{.j}}$
x_2	$\frac{p_{2j}}{p_{.j}}$
$\vdots$	$\vdots$
x_n	$\frac{p_{nj}}{p_{.j}}$

توزیع‌های احتمال شرطی Y به شرط $X = x_i$

Y	y_1	y_2	…	y_m
$P(Y \vert X = x_i)$	$\frac{p_{i1}}{p_{i.}}$	$\frac{p_{i2}}{p_{i.}}$		$\frac{p_{im}}{p_{i.}}$

توزیع‌های دو متغیره پیوسته

درحالت پیوسته، توزیع مشترک دو متغیر تصادفی X و Y به صورت تابع چگالی مشترک بیان می‌شود.

تابع چگالی مشترک دو متغیر تصادفی X و Y به صورت $f(x,y)$ بیان می‌شود.

مجموع احتمالات برابر با ۱ است (شرط نرمال بودن به صورت زیر باید برقرار باشد تا مطمئن شویم که این تابع، یک تابع چگالی احتمال معتبر است):

$$\int_y \int_x f(x,y)dxdy = 1$$

تابع چگالی حاشیه‌ای X

$$f(x) = \int_y f(x,y)dy$$

تابع چگالی حاشیه‌ای Y

$$f(y) = \int_x f(x,y)dx$$

توزیع‌های دو متغیره پیوسته

تابع چگالی مشترک دو متغیر تصادفی X و Y به صورت $f(x,y)$

تابع توزیع شرطی X

$$f(x|y) = \frac{f(x,y)}{f(y)}$$

تابع توزیع شرطی Y

$$f(y|x) = \frac{f(x,y)}{f(x)}$$

توزیع‌های دو متغیره پیوسته

مثال

تابع چگالی مشترک X و Y به صورت زیر درنظر بگیرید.

$$f(x,y)=\frac{3}{5}x(y+x) \quad ; 0<x<1 \ , 0<y<2$$

مطلوبست توابع چگالی حاشیه‌ای و شرطی X و Y ؟

توزیع‌های دو متغیره پیوسته

مثال

تابع چگالی حاشیه‌ای X

$$f(x) = \int_0^2 \frac{3}{5}x(y+x)dy = \frac{3}{5}x\left(\frac{y^2}{2}+xy\right) = \frac{6}{5}x(1+x)$$

تابع چگالی حاشیه‌ای Y

$$f(y) = \int_0^1 \frac{3}{5}x(y+x)dx = \frac{3}{5}\left(\frac{y}{2}+\frac{1}{3}\right)$$

توزیع‌های دو متغیره پیوسته

مثال

تابع توزیع شرطی X

$$f(x|y) = \frac{\frac{3}{5}x(y+x)}{\frac{3}{5}\left(\frac{y}{2}+\frac{1}{3}\right)} = \frac{6x(y+x)}{2+3y}$$

تابع توزیع شرطی Y

$$f(y|x) = \frac{(y+x)}{2(1+x)}$$

همبستگی دو متغیر X و Y

برای بررسی همبستگی دو متغیر X و Y معمولاً از کوواریانس استفاده می‌شود؛

$$\operatorname{cov}(X,Y)=E(XY)-E(X).E(Y)$$

در حالت پیوسته؛

$E(X)$: با استفاده از تابع چگالی حاشیه‌ای X محاسبه می‌شود: $E(X)=\int xf(x)dx$

$E(Y)$: با استفاده از تابع چگالی حاشیه‌ای Y محاسبه می‌شود: $E(Y)=\int yf(y)dy$

$$E(XY)=\int_y\int_x xyf(x,y)dxdy$$

امید ریاضی شرطی دو متغیر X و Y

یکی از معیارهای مهم برای ارزیابی رابطه دو متغیر تصادفی، امید ریاضی شرطی است.

امید ریاضی شرطی را می‌توان با استفاده از توابع چگالی شرطی به دست آورد؛

$$E(X|y) = \int_x xf(x|y)dx = h(y)$$

$$E(Y|x) = \int_y yf(y|x)dy = g(x)$$

اگر امید ریاضی شرطی برای یک متغیر تابعی از متغیر دیگر باشد، آنگاه این دو متغیر همبستگی دارند.

از طرف دیگر، اگر امیدریاضی شرطی Y تابعی از X نباشد، آنگاه X و Y مستقل از هم هستند. در این صورت، امیدریاضی شرطی و غیرشرطی با هم برابرند: $E(Y|x) = E(Y)$