Gamma函数以及分布的理解

1：数学公式

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} \mathrm{dt}$$

2：数学性质

$$\Gamma(x+1) = x\Gamma(x)$$

3：推论

$$\Gamma(n+1) = n\Gamma(n-1) = \ldots = n!$$

Gamma分布

1：Gamma函数到Gamma分布过程

Gamma函数同时除以$\Gamma(x)$，公式变化过程如下：

$$1 = \int_0^\infty \frac{t^{x-1} e^{-t}}{\Gamma(x)} \mathrm{dt}$$

进行符号转化，t换成x，x换成α：

$$1 = \int_0^\infty \frac{x^{\alpha-1} e^{-x}}{\Gamma(\alpha)} \mathrm{dx} \tag{1}$$

令$x = \beta t$

$$1 = \int_0^\infty \frac{(\beta t)^{\alpha-1} e^{-\beta t}}{\Gamma(\alpha)} \mathrm{d}\beta \mathrm{t}$$

$$1 = \int_0^\infty \frac{\beta^\alpha t^{\alpha-1} e^{-\beta t}}{\Gamma(\alpha)} \mathrm{dt} \tag{2}$$

若是按照上述式子1，则得到如下概率密度的gamma分布：

$$f(x|\alpha) = \frac{x^{\alpha-1} e^{-x}}{\Gamma(\alpha)} \tag{3}$$

若是按照上述式子2，则得到如下概率密度的gamma分布：

$$f(x|\alpha, \beta) = \frac{\beta^\alpha x^{\alpha-1} e^{-\beta x}}{\Gamma(\alpha)} \tag{4}$$

两种Gamma分布的形式，其中式子4为其一般的定义式子，分布的取决于两个参数，当$\beta=1$时，可以看到式子4就退化到式子3了，式子3被称为标准gamma分布。

2：Gamma分布的性质

1）当$\beta=1,\alpha=1$时gamma分布退化为标准指数分布

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

2)期望值:

$$\begin{aligned}
E(f(x)) &= \int_0^\infty xf(x)dx \\
&= \int_0^\infty x\frac{\beta^\alpha x^{\alpha-1}e^{-\beta x}}{\Gamma(\alpha)}dx \\
&= \frac{\beta^\alpha}{\Gamma(\alpha)}\int_0^\infty x^\alpha e^{-\beta x}\Gamma(\alpha)dx \\
&= \frac{\beta^\alpha}{\Gamma(\alpha)}\int_0^\infty (t/\beta)^\alpha e^{-\beta(t/\beta)}d(t/\beta) \\
&= \frac{\beta^\alpha}{\beta\Gamma(\alpha)\beta^\alpha}\int_0^\infty (t)^\alpha e^{-t}dt \\
&= \frac{\beta^\alpha}{\beta\Gamma(\alpha)\beta^\alpha}\Gamma(\alpha+1) \\
&= \frac{\Gamma(\alpha+1)}{\beta\Gamma(\alpha)} \\
&= \frac{\alpha\Gamma(\alpha)}{\beta\Gamma(\alpha)} \\
&= \frac{\alpha}{\beta}
\end{aligned}$$

2)方差值:

$$D(f(x))=\frac{\alpha}{\beta^2}$$

3：Gamma分布到泊松分布

3.1：泊松分布

1）泊松分布定义

$$P(X=k|\lambda)=\frac{\lambda^k e^{-\lambda}}{k!}$$

2）gamma分布，当$\beta=1$时，标准gamma分布:

$$f(x|\alpha=k+1)=\frac{x^k e^{-x}}{\Gamma(k+1)}=\frac{x^k e^{-x}}{k!}$$

对比看是泊松分布对应这离散随机量，gamma分布对应这连续的量，数学形式一样，变动的量不同而已。

3.2：二项分布

1)二项分布定义如下：

$$P(X=k|n,p)=C_n^k p^k(1-p)^{n-k}$$

其中p表示单次实验的成功的概率，n表示实验的次数，k表示成功的次数

2)二项分布的累计概率分布如下：

$$P(X\leq k)=\frac{n!}{k!(n-k-1)!}\int_p^1 t^k(1-t)^{n-k-1}dt \tag{5}$$

证明过程如下：

$$\begin{aligned}\frac{n!}{k!(n-k-1)!}\int_p^1 t^k(1-t)^{n-k-1}dt &= \frac{n!}{k!(n-k-1)!}\{-\frac{kt^{k-1}(1-t)^{n-k}}{n-k}|_p^1+\int_p^1\frac{kt^{k-1}(1-t)^{n-k}}{n-k}\}\\ &= \frac{n!}{(k-1)!(n-k)!}\{p^{k-1}(1-p)^{n-k}+\int_p^1 t^{k-1}(1-t)^{n-k}dt\}\\ &= C_n^k p^k(1-p)^{n-k}+\frac{n!}{(k-1)!(n-k)!}\int_p^1 t^{k-1}(1-t)^{n-k}dt\\ &= P(X=k|n,p)+P(X\leq k-1)\end{aligned}$$

根据上面的推导可得如下，最终能够说明上述式子成立：

$$P(X\leq k)=P(X=k|n,p)+P(X\leq k-1)$$

3.3：二项分布，泊松分布，Gamma分布关联

1）二项分布到泊松分布过程

令$np=\lambda$，则$p=\frac{\lambda}{n}$,下面推导利用定理$\lim_{x\to+\infty}(1+\frac{a}{x})^x=e^a$这个定理也会好证明，只需要取$log$后，采用洛必达法则即可证明。

$$\begin{aligned}\lim_{n\to\infty}P(X=k|n,p) &= \lim_{n\to\infty}\frac{n!}{k!(n-k)!}(\frac{\lambda}{n})^k(1-\frac{\lambda}{n})^{n-k}\\ &= \lim_{n\to\infty}\frac{\lambda^k}{k!}\frac{n!}{(n-k)!n^k}(1-\frac{\lambda}{n})^n(1-\frac{\lambda}{n})^{-k}\\ &= \lim_{n\to\infty}\frac{\lambda^k}{k!}\frac{n(n-1)\ldots(n-k+2)(n-(k-1))}{n^k}(1-\frac{\lambda}{n})^n(1-\frac{\lambda}{n})^{-k}\\ &= \lim_{n\to\infty}\frac{\lambda^k}{k!}[(1-\frac{k-1}{n})(1-\frac{k-2}{n})\ldots(1-\frac{1}{n})](1-\frac{\lambda}{n})^n(1-\frac{\lambda}{n})^{-k}\\ &= \lim_{n\to\infty}\frac{\lambda^k}{k!}(1-\frac{\lambda}{n})^n\\ &= \frac{\lambda^k}{k!}\lim_{n\to\infty}(1-\frac{\lambda}{n})^n=\frac{\lambda^k}{k!}e^{-\lambda}\end{aligned}$$

这里表明，当二项分布做无穷次实验时，会让其变为泊松分布。

2）基于式子5，令$nt=x$,则$t=\frac{x}{n}$可以得到如下不等式

$$\begin{aligned}P(X\leq k|n,p) &= \frac{n!}{k!(n-k-1)!}\int_{np}^n(\frac{x}{n})^k(1-\frac{x}{n})^{n-k-1}d\frac{x}{n}\\ &= \frac{(n-1)!}{k!(n-k-1)!}\int_{np}^n(\frac{x}{n})^k(1-\frac{x}{n})^{n-k-1}dx\\ &= \int_{np}^n P(X=k|n-1,\frac{x}{n})dx\end{aligned}$$

基于上述1）的推论，当$n \to +\infty$时，则转为了泊松分布，其中对于二项分$B(X|n, p)$转为了$Poisson(X|\lambda = np)$,对于$B(X|n-1, \frac{x}{n})$转为$Poisson(X|\lambda = \lim_{n\to\infty}(n-1)\frac{x}{n} = x)$，从而转为了泊松的如下形式：

$$Poisson(X \leq k|\lambda) = \int_{\lambda}^{+\infty} Poisson(X = k|x)dx \tag{6}$$

特别的情况下，当$\lambda \to 0$时，泊松分布的概率性趋向于确定性，也就是说当$\lambda = 0$时，这样的事件就是确定性事件了，也就是不管k值如何取，都不可能发生的事件了，从而确定性事件分布，从而其概率为1。那么会得到如下式子

$$1 = \int_0^{+\infty} \frac{e^{-x}x^k}{k!}dx$$

根据gamma的分布式子1，当$\alpha = k + 1$

$$1 = \int_0^{\infty} \frac{x^k e^{-x}}{\Gamma(k+1)} \mathrm{dx} = \int_0^{\infty} \frac{x^k e^{-x}}{k!} dx$$

从上述可以看到gamma分布和泊松分布的关系
由式子6可以附带有意思的推论式子：

$$Poisson(X \leq k|\lambda) + \int_0^{\lambda} Poisson(X = k|x)dx = 1$$

$$Poisson(X \leq k|\lambda) + \int_0^{\lambda} \frac{e^{-x}x^k}{k!}dx = 1$$

$$Poisson(X \leq k|\lambda) + \int_0^{\lambda} Gamma(x|\alpha = k+1, \beta = 1)dx = 1 \tag{7}$$

式子7表明了泊松分布和gamma的关系，**gamma的积分与泊松的累计概率求和为1**。

> **参考资料:**
> LDA数学八卦
> gamma, 二项式, 泊松分布的百度百科