جلسه‌ی پانزدهم آمار و احتمال مهندسی «دکتر شریفی زارچی» - دانشگاه شریف

مجموع متغیرهای تصادفی

اگر X,Y از توزیع پواسون باشند؛ Z از چه توزیعی است؟

$X \sim Poisson(\lambda_1)\ ,\ Y \sim Poisson(\lambda_2)$

if $Z = X + Y$ then $Z \sim ?$

مثل مجموع بینندگان یک فیلم = بینندگان آقا + بینندگان خانم

یا مثلاً لود روی سرور بانک = تراکنش‌های مانده حساب + تراکنش‌های کارت به کارت + تراکنش‌های خرید شارژ + ...

$$E(Z) = \sum_i E(X_i) \quad , \quad Var(Z) = \sum_i Var(X_i) \quad , \quad Z = \text{?}$$

یعنی از چه توزیعی پیروی میکند.

کاربرد همچین توزیع: $\mathbb{P}(Z > \alpha)$ مثلاً احتمال اینکه مقدار تراکنش‌های روی سرور بانک از یه حدی بالاتر بره سرور به مشکل میخوره. ما باید بفهمیم احتمال این میزان تراکنش چقدره و این اهمیت همچین توزیع Z است.

$$\mathbb{P}(Z=z) = \sum_{k=0}^{z} \mathbb{P}(X=k)\,\mathbb{P}(Y=z-k) = \sum_{k=0}^{z} \frac{e^{-\lambda_1}\lambda_1^k}{k!}\,\frac{e^{-\lambda_2}\lambda_2^{(z-k)}}{(z-k)!}$$

$$= \frac{e^{-(\lambda_1+\lambda_2)}}{z!} \sum_{k=0}^{z} \binom{z}{k} \lambda_1^k \lambda_2^{z-k} = \frac{e^{-(\lambda_1+\lambda_2)}}{z!} (\lambda_1+\lambda_2)^z$$

و درنتیجه داریم: if $X \sim Poisson(\lambda_1)$, $Y \sim Poisson(\lambda_2)$ then: $Z \sim Poisson(\lambda_1 + \lambda_2)$

نکته حاصل جمع دو متغیر تصادفی از یک توزیع یکسان (هر دو از یک توزیع باشند)؛ از همان توزیع پیروی خواهد کرد. (با شرط مستقل بودن)

$X, Y \sim Uniform(0,1)$, X,Y مستقلند.

$Z = X + Y,\ Z = ?$

$$F_Z(z) = \mathbb{P}(Z \le z) = \iint_{x+y \le z} f_{XY}(x,y)\,dx\,dy$$

$$= \iint_{x+y\le z} f_X(x) f_Y(y)\,dx\,dy = \int_{-\infty}^{+\infty} \Big(\underbrace{\int_{-\infty}^{z-y} f_X(x)\,dx}_{F_X(z-y)} \Big) f_Y(y)\,dy = \int_{-\infty}^{+\infty} F_X(z-y) f_Y(y)\,dy$$

$$f_Z(z) = \frac{d}{dz} \int_{-\infty}^{+\infty} F_X(z-y)\, f_Y(y)\, dy = \int_{-\infty}^{+\infty} f_X(z-y)\, f_Y\, dy$$

نکته

$$\Rightarrow f_Z(z) = \underbrace{(f_X * f_Y)}_{\text{Convolution}}(z)$$

به ازای هر X, Y که مستقل و پیوسته باشند برقرار است.

و حالا اگه تعداد متغیرهای تصادفی که جمع می‌زنیم رو زیاد کنیم. $X_1 + X_2 + \dots + X_n$ و همه اینا دو به دو مستقل باشند و از یک توزیع یکسان پیروی کنند در نهایت $Z = \sum_i X_i$ ها از توزیع نورمال پیروی خواهند کرد. هر چقدر n بزرگتر باشه. توزیع ما نرمال‌تر میشه.

برای جمع چندتا متغیر تصادفی، دوتای ابتدایی رو Convolution می‌گیریم و حاصل رو با متغیر سوم Convolution می‌کنیم و به همین ترتیب تا انتها.

$$X, Y \sim N(0,1)$$

$$Z = aX + bY \sim\ ?$$

نرمال با واریانس a^2+b^2

اثبات خیلی سنگینی داره که توی جزوه گیت هاب هست و نوشتن نداره :)