Lecture Stat 302
Introduction to Probability - Slides 19

AD

April 2010

Sum of Independent Random Variables

- Consider $Z = X + Y$ where X and Y are disrete r.v. of respective p.m.f. $p_X(x)$ and $p_Y(y)$ then

$$p_Z(z) = \sum_y p_X(z - y)\, p_Y(y).$$

- Consider $Z = X + Y$ where X and Y are continuous r.v. of respective p.d.f. $f_X(x)$ and $f_Y(y)$ then

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(z - y)\, f_Y(y)\, dy.$$

Sum of Exponential Random Variables

- Consider two independent exponential r.v. X, Y of parameter λ (i.e. $f_X(x) = f_Y(x) = \lambda e^{-\lambda x} 1_{[0,\infty)}(x)$).
- The pdf of $Z = X + Y$ is $f_Z(z) = 0$ for $z < 0$ and for $z > 0$

$$\begin{aligned}
f_Z(z) &= \int_{-\infty}^{\infty} \lambda e^{-\lambda(z-y)} 1_{[0,\infty)}(z-y) \; \lambda e^{-\lambda y} 1_{[0,\infty)}(y) \, dy \\
&= \lambda^2 e^{-\lambda z} \int_0^{\infty} 1_{[0,\infty)}(z-y) \, dy \\
&= \lambda^2 e^{-\lambda z} \int_0^{z} 1_{[0,\infty)}(z-y) \, dy \\
&= \lambda^2 z e^{-\lambda z}.
\end{aligned}$$

Sum of Gaussian Random Variables

- Consider two independent normal standard r.v. X, Y (i.e. $f_X(x) = f_Y(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$) then the pdf of $Z = X + Y$ is

$$f_Z(z) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-(z-y)^2/2} e^{-y^2/2} dy$$

where

$$\begin{aligned} (z-y)^2 + y^2 &= z^2 + 2y^2 - 2yz \\ &= z^2 + 2(y - z/2)^2 - z^2/2 = 2(y - z/2)^2 + z^2/2 \end{aligned}$$

- So we have

$$f_Z(z) = \frac{e^{-z^2/4}}{2\pi} \sqrt{\pi} \int_{-\infty}^{\infty} \frac{1}{\sqrt{\pi}} e^{-(y-z/2)^2} dy = \frac{e^{-z^2/4}}{\sqrt{2\pi}\sqrt{2}}$$

Hence Z is a normal r.v. of mean 0 and variance 2.

- **Generalization**: if X is an normal r.v. (μ_X, σ_X^2) and Y is an normal r.v. (μ_Y, σ_Y^2) where X and Y are independent then Z is a normal r.v. $(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2)$.

Conditional Distributions: Discrete Case

- Given a joint p.m.f. for two r.v. X, Y it is possible to compute the conditional p.m.f. X given $Y = y$.
- Assume X, Y are discrete-valued r.v. with a joint p.m.f. $p(x, y)$ then the conditional p.m.f. of X given $Y = y$ is

$$\begin{aligned} p_{X|Y}(x|y) &:= P(X = x|Y = y) \\ &= \frac{P(X = x \cap Y = y)}{P(Y = y)} \\ &= \frac{p(x, y)}{p_Y(y)}. \end{aligned}$$

- In the case where X and Y are independent, we have $p_{X|Y}(x|y) = p_X(x)$ as $p(x, y) = p_X(x)\, p_Y(y)$.

Conditional Distributions: Discrete Case

- We have

$$p(x, y) = p_{X|Y}(x|y)\, p_Y(y)$$

and similarly

$$p(x, y) = p_{Y|X}(y|x)\, p_X(x)$$

- Hence we obtain

$$p_{X|Y}(x|y) = \frac{p_{Y|X}(y|x)\, p_X(x)}{p_Y(y)}$$

which holds if $p_Y(y) > 0$.

Conditional Expectation and Variance: Discrete Case

- We can define the mean, variance of the conditional p.m.f.
- The conditional mean is given by

$$E(X|Y=y)=\sum_x x\cdot p_{X|Y}(x|y)$$

- The conditional variance is given by

$$\begin{aligned} Var(X|Y=y) &= E\left((X-E(X|Y=y))^2\,|Y=y\right) \\ &= E(X^2|Y=y)-\{E(X|Y=y)\}^2 \end{aligned}$$

where

$$E(X^2|Y=y)=\sum_x x^2\cdot p_{X|Y}(x|y)$$

- $E(X|Y=y)$ and $Var(X|Y=y)$ are functions but $E(X|Y)$ and $Var(X|Y)$ are random variables.

Example: Toy problem

- Consider $X \in \{0,1,2\}$ and $Y \in \{0,1,2\}$ such that their joint pmf is given by

		y		
		0	1	2
x	0	1/9	2/9	1/9
	1	2/9	2/9	0
	2	1/9	0	0

- The conditional pmf of X given $Y=0$ and $Y=1$ are

$$p_{X|Y}(x|\,0) = \frac{p(x,y=0)}{1/9+2/9+1/9},\; p_{X|Y}(x|\,1) = \frac{p(x,y=1)}{2/9+2/9+0}.$$

- We have

$$\begin{aligned} E(X|\,0) &= 1 \times p_{X|Y}(x|\,0) + 2 \times p_{X|Y}(x|\,0) \\ &= \frac{2/9}{4/9} + 2 \times \frac{1/9}{4/9} = 1 \end{aligned}$$

Example: Fair Die

- Roll a die until we get a 6. Let Y be the total number of rolls and X the number of 1's we get. What is the conditional pmf $p_{X|Y}(x|y)$? Compute $E(X|y)$ and $Var(X|Y=y)$.
- The event $Y=y$ means that there were $y-1$ rolls that were not a 6 and then the yth roll was a six.
- So $p_{X|Y}(x|y)$ is a binomial distribution with $n=y-1$ trials and proba. of success $p=1/5$.
- It follows that

$$\begin{aligned} E(X|y) &= np = \frac{(y-1)}{5}, \\ Var(X|y) &= np(1-p) = \frac{4(y-1)}{25} \end{aligned}$$

- If we do not observe $Y=y$, then $E(X|Y) = \frac{Y-1}{5}$ and $Var(X|Y) = \frac{4(Y-1)}{25}$ are not numbers but random variables.

Example: Fair Die

- As $E(X|Y)$ is a random variable, it is possible to compute its expectation

$$\begin{aligned} E(E(X|Y)) &= \sum_y E(X|Y=y).p_Y(y) \\ &= \sum_y \frac{(y-1)}{5}.p_Y(y) \\ &= -\frac{1}{5} + \sum_y y p_Y(y) = -\frac{1}{5} + E(Y) \\ &= -\frac{1}{5} + 6 \text{ as } Y \text{ Geometric} \end{aligned}$$

- It can actually be easily established that

$$E(E(X|Y)) = E(X).$$

Example: Poisson random variables

- Consider two Poisson independent r.v. X and Y of respective p.m.f. $p_X(x) = e^{-\lambda} \frac{\lambda^x}{x!}$ for $x = 0, 1, 2, ...$ and $p_Y(y) = e^{-\lambda'} \frac{\lambda'^y}{y!}$. Calculate the conditional p.m.f. of X given $X + Y = m$. What is the conditional expectation and variance $E(X|y)$ and $Var(X|y)$?
- We have $p(X = x|X + Y = m) = 0$ for $x > m$ and for $x \leq m$

$$
\begin{aligned}
p(X = x|X + Y = m) &= \frac{p(X=x, X+Y=m)}{p(X+Y=m)} \\
&= \frac{p(X=x, Y=m-x)}{p(X+Y=m)} = \frac{p_X(X=x)p_Y(Y=m-x)}{p(X+Y=m)} \\
&= \frac{e^{-\lambda} \frac{\lambda^x}{x!} e^{-\lambda'} \frac{(\lambda')^{m-x}}{(m-x)!}}{e^{-(\lambda+\lambda')} \frac{(\lambda+\lambda')^m}{m!}} \text{ as } X + Y \text{ is Poisson } \lambda + \lambda' \\
&= \binom{m}{x} \frac{\lambda^x (\lambda')^{m-x}}{(\lambda+\lambda')^m} = \binom{m}{x} \left(\frac{\lambda}{\lambda+\lambda'}\right)^x \left(1 - \frac{\lambda}{\lambda+\lambda'}\right)^{m-x}
\end{aligned}
$$

which is a Binomial of parameter $n = m$ and success proba $p = \lambda / (\lambda + \lambda')$.

Example: Poisson random variables

- Hence we have

$$E\left(X \mid X+Y=m\right)=np=\frac{m\lambda}{\left(\lambda+\lambda'\right)}$$

- We also obtain

$$\begin{aligned} Var\left(X \mid X+Y=m\right) &= np\left(1-p\right) \\ &= \frac{m\lambda\lambda'}{\left(\lambda+\lambda'\right)^2} \end{aligned}$$

Example: How Many Tax Fraudsters?

- The number N of tax fraudsters is assumed to follow a Poisson distribution with param λ. Each tax fraudster is identified with proba p independently of the other fraudsters. Let K be the number of fraudsters identified. What is the conditional p.m.f. of N given $K = k$? Compute $E\left(N\middle|\, k\right)$.
- We have

$$\begin{aligned} p_N(n) &= e^{-\lambda}\frac{\lambda^n}{n!}, \\ p_{K|N}(k|\, n) &= \binom{n}{k} p^k (1-p)^{n-k}. \end{aligned}$$

- We want to compute

$$p_{N|K}(n|\, k) = \frac{p_{K|N}(k|\, n)\, p_N(n)}{p_K(k)}$$

Example: How Many Tax Fraudsters?

- We have $p_{N|K}\left(n|\,k\right)=0$ if $n<k$.
- If $n\geq k$

$$
\begin{aligned}
p_{N|K}\left(n|\,k\right) &= \frac{\begin{pmatrix} n \\ k \end{pmatrix} p^k\left(1-p\right)^{n-k} e^{-\lambda}\frac{\lambda^n}{n!}}{\sum_{m\geq k}\begin{pmatrix} m \\ k \end{pmatrix} p^k\left(1-p\right)^{m-k} e^{-\lambda}\frac{\lambda^m}{m!}} \\
&= \frac{\left\{\left(1-p\right)\lambda\right\}^{n-k}}{\left(n-k\right)!}e^{-\left(1-p\right)\lambda} \text{ (use } i=m-k\text{)}
\end{aligned}
$$

- Hence we have

$$
E\left(N|\,k\right)=\sum_{n\geq k} n\frac{\left\{\left(1-p\right)\lambda\right\}^{n-k}}{\left(n-k\right)!}e^{-\left(1-p\right)\lambda}=k+\left(1-p\right)\lambda
$$