Lecture 20: Multinomial and Cauchy

Example find $E|Z_1 - Z_2|$, with $Z_1, Z_2 \overset{\text{i.i.d.}}{\sim} N(0,1)$.

Note: $Z_1 - Z_2 \sim N(0, 2)$ add the variances

So, $E|Z_1 - Z_2| = E|\sqrt{2}Z|$, $Z \sim N(0,1)$ (Standard Normal)

$= \sqrt{2}\, E|Z|$

$= \sqrt{2}\int_{-\infty}^{\infty} \underbrace{|z| \frac{1}{\sqrt{2\pi}} e^{-z^2/2}}_{\text{even function}} dz$ (By LOTUS 1-D)

$= 2\sqrt{2}\int_0^{\infty} z \times \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz$

$= \frac{2\sqrt{2}}{\sqrt{2\pi}}\int_0^{\infty} z e^{-z^2/2} dz$

$= \frac{2\cancel{\sqrt{2}}}{\cancel{\sqrt{2}}\sqrt{\pi}} = 2/\sqrt{\pi}$

$u = \int_0^{\infty} z e^{-z^2/2} dz$

let $+z^2/2 = v$

$+\frac{\cancel{2}z\,dz}{\cancel{2}} = dv$

$z\,dz = +dv$

$u = \int_0^{\infty} e^{-v} dv$

$u = -[e^{-v}]_0^{\infty}$

$u = -[e^{-\infty} - e^0]$

$= -[\cancel{e^{-\infty}}^{0}$

$= e^0 - \cancel{1/e^{\infty}}$

$= 1$

Theorem: $X \sim N(\mu_1, \sigma_1^2)$, $Y \sim N(\mu_2, \sigma_2^2)$

independent. Then, $X+Y \sim N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$.

Proof: Use MGFs:

MGF of $X+Y$ is $\underbrace{e^{\mu_1 t + 1/2 \sigma_1^2 t^2}}_{MGF(X)} \cdot \underbrace{e^{\mu_2 t + 1/2 \sigma_2^2 t^2}}_{MGF(Y)}$

$$= \underbrace{e^{(\mu_1+\mu_2)t + 1/2(\sigma_1^2+\sigma_2^2)t^2}}_{\text{MGF of } N(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)}$$

~~Multinomial Distribution~~

→ A multivariate distribution means a joint distribution for more than one random variables.

Univariate distributions (one random variable)

Normal

Poisson

Geometric

Multivariate distributions

Multinomial (generalization of binomial to higher dimension)

multivariate normal (generalization of normal distribution to higher ~~distributio~~ dimensions.)

Multinomial distribution

→ Higher dimensional version of the binomial distribution.

Definition/Story of $Mult(n, \vec{p})$, $\vec{p} = (p_1, \dots, p_K)$ probability vector

$$p_j \geq 0, \sum_j p_j = 1$$

In binomial, there are two possible outcomes, i.e., two categories (success and failure).
Multinomial means instead of two categories, we have K categories.

$$\vec{X} \sim Mult(n, \vec{p}), \quad \vec{X} = (X_1, \dots, X_K)$$

If we have n objects, which we are independently putting into K categories,

$p_j = P(\text{category } j)$ → Any of these n objects is in category j, has probability p_j

$X_j = \#$ objects in category j.

Joint PMF

$$P(X_1 = n_1, \dots, X_K = n_K) = \frac{n!}{n_1! n_2! \dots n_K!} p_1^{n_1} p_2^{n_2} \dots p_K^{n_K}$$

→ if $n_1 + n_2 + \dots + n_K = n$

(0, otherwise)

(what's the prob. that there are n1 objects in the first category, n2 in second and so on)

Example

$\vec{X} \sim Mult_k(n, \vec{p})$. $k \rightarrow$ no. of categories | dimensions

Find marginal distribution of X_j (how many objects are in categories j).

Then $X_j \sim Bin(n, p_j)$
(each of these objects are either in category j or isn't)
we are assuming they are all independent trials.

$E(X_j) = np_j$, $Var(X_j) = np_j(1-p_j)$

Lumping property

$\vec{X} = (X_1, X_2, \ldots, X_{10}) \sim Mult(n, (p_1, \ldots, p_{10}))$ prop. of diff. party memberships

Example Let's imagine we are in a country that has 10 political parties. We take n people and assume that the people are independent of each other and everyone in this country is a member of one of these 10 parties.
X_1 is the number of people in the first political party,
X_2 is the number in second one, and so on.

→ So, what if it's a country where there are 2 dominant parties and all the other parties are much smaller. So, it might be kind of unwieldy to deal with above 10 dimensional vector. Suppose that the first two are ~~kind of~~ the two dominant major parties, the rest of them are minor, so we may wanna just lump them together.

Hence, the lumping property, i.e., lump all the other parties together.

Let $\vec{Y} = (X_1, X_2, X_3 + X_4 + \dots + X_{10})$. Then

$$\vec{Y} \sim \text{Mult}(n, (p_1, p_2, p_3 + \dots + p_{10}))$$

(we need to make sure that each object is in exactly one category. so, it wouldn't work if we could be in more than one category or be in no categories.)

Conditional distribution

$\vec{X} \sim \text{Mult}(n, \vec{p})$. Then given $X_1 = n_1$,

$$(X_2, \dots, X_k) \sim \text{Mult}_{k-1}(n - n_1, (p_2', \dots, p_n')),$$

with $p_2' = P(\text{being in category } 2 \mid \text{not in category } 1)$

$$= \frac{p_2}{1 - p_1} = \frac{p_2}{p_2 + \dots + p_k}$$

$$p_j' = \frac{p_j}{p_2 + \dots + p_k}$$

Example **Cauchy Interview Problem**

The Cauchy ~~distribution~~ is the distribution of $T = X/Y$ with $X, Y \overset{\text{i.i.d.}}{\sim} N(0,1)$. Find PDF of T.

Solu: $\underbrace{P\left(X/Y \leq t\right)}_{\text{event}} = P\left(\frac{X}{|Y|} \leq t\right)$ this follows from the symmetry of Normal i.e., $N(0,1)$

CDF $F(t)$

$$= P(X \leq t|Y|)$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{t|Y|} \underbrace{\frac{1}{\sqrt{2\pi}} e^{-x^2/2}}_{PDF(x)} \underbrace{\frac{1}{\sqrt{2\pi}} e^{-y^2/2}}_{PDF(y)} dx\, dy$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-y^2/2} \int_{-\infty}^{t|Y|} \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx\, dy$$

$$= \frac{1}{\sqrt{2\pi}} \underbrace{\int_{-\infty}^{\infty} e^{-y^2/2} \Phi(t|Y|)\, dy}_{\text{even function}}$$

$$= \sqrt{2/\pi} \int_{0}^{\infty} e^{-y^2/2} \Phi(ty)\, dy$$

PDF: $F'(t) = \sqrt{2/\pi} \int_0^\infty e^{-y^2/2} \left(y \frac{1}{\sqrt{2\pi}} e^{-t^2y^2/2} dy\right)$

(derivation of standard normal CDF is the standard normal PDF)

$$= \frac{1}{\pi} \int_0^\infty y\, e^{-(1+t^2)y^2/2} dy = \frac{1}{\pi(1+t^2)}$$

let $u = (1+t^2)y^2/2$

$\Rightarrow du = \frac{2y}{2}(1+t^2)\, dy \Rightarrow du = y(1+t^2)\, dy$

Alternate method
(Law of total prob.)

$$P(X \leq t|Y|) = \int_{-\infty}^{\infty} P(X \leq t|Y| \mid Y = y) \underbrace{\phi(y)}^{N(0,1)\,PDF} dy$$

(Conditioned on Y)

X, Y independent

$$= \int_{-\infty}^{\infty} \Phi(t|y|)\phi(y)\,dy$$

$t y$ is some constant

$P(X \leq$ some constant $t y)$

It's just the standard normal CDF evaluated there.