" $\Rightarrow P[\underset{\sim}{X} \in \underbrace{(\underset{\sim}{x}, \underset{\sim}{x}+d\underset{\sim}{x}]}_{(x_1, x_1+dx_1] \times (x_2; x_2+dx_2] \times \cdots \times (x_k, x_k+dx_k]}] = f_{\underset{\sim}{X}}(\underset{\sim}{x})\, d\underset{\sim}{x}$,

i.e.,

$P[x_1 < X_1 < x_1+dx_1,\ x_2 < X_2 < x_2+dx_2, \ldots, x_k < X_k \leq x_k+dx_k]$

$= f_{X_1, X_2, \ldots, X_k}(x_1, x_2, \ldots, x_k)\, dx_1\, dx_2 \cdots dx_k$ "

Any joint pdf $f_{\underset{\sim}{X}}$ satisfies

(I) $f_{\underset{\sim}{X}}(\underset{\sim}{x}) \geq 0 \quad \forall \quad \underset{\sim}{x} \in \mathbb{R}^k$, and

(II) $\int \cdots \int_{\mathbb{R}^k} f_{\underset{\sim}{X}}(\underset{\sim}{x})\, d\underset{\sim}{x} = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} f_{X_1, \ldots, X_k}(x_1, \ldots, x_k)\, dx_1 \ldots dx_k = 1.$

Again from the joint cdf $F_{\underset{\sim}{X}}(\underset{\sim}{x})$, a joint pdf can be guessed using the following recipe:

$$f_{\underset{\sim}{X}}(\underset{\sim}{x}) = \begin{cases} \dfrac{\partial}{\partial x_1} \dfrac{\partial}{\partial x_2} \cdots \dfrac{\partial}{\partial x_k} F_{\underset{\sim}{X}}(\underset{\sim}{x}) & \text{whenever the partial derivatives exist,} \\ 0 & \text{otherwise.} \end{cases}$$

If $\underset{\sim}{X}$ is indeed a cont random vector, then the above recipe will always work and the order of the partial derivatives ~~does~~ will not matter.

Independence of k random variables

Suppose $\underset{\sim}{X} = (X_1, X_2, \ldots, X_k)$ is any (not necessarily discrete or continuous) random vector.

Defn: We say that the r.v.s $X_1, X_2, \ldots, X_k$ are independent if $\forall\ \underset{\sim}{u} = (u_1, u_2, \ldots, u_k) \in \mathbb{R}^k$,

$$P(X_1 \leq u_1, X_2 \leq u_2, \ldots, X_k \leq u_k) = \prod_{i=1}^{k} P(X_i \leq u_i),$$

i.e., $F_{\underset{\sim}{X}}(\underset{\sim}{u}) = F_{X_1}(u_1) F_{X_2}(u_2) \ldots F_{X_k}(u_k)$.

Roughly speaking, this means that the r.v.s $X_1, X_2, \ldots, X_k$ do not influence each other.

In parallel to the bivariate case, the following theorems can be proved.

ThmD: Suppose $\underset{\sim}{X}$ is a discrete random vector. Then $X_1, X_2, \ldots, X_k$ are independent if and only if $p_{\underset{\sim}{X}}(\underset{\sim}{x}) = p_{X_1}(x_1)\, p_{X_2}(x_2) \ldots p_{X_k}(x_k)\ \forall\ \underset{\sim}{x} \in \mathbb{R}^k$.

Thm C: Suppose $X_1, X_2, \dots, X_k$ are cont r.v.s with pdfs $f_{X_1}, f_{X_2}, \dots, f_{X_k}$, respectively. Then $X_1, X_2, \dots, X_k$ are independent if and only if $\underset{\sim}{X}$ is a cont random vector with a joint pdf

$$h(\underset{\sim}{x}) = f_{X_1}(x_1) f_{X_2}(x_2) \dots f_{X_k}(x_k), \quad \underset{\sim}{x} \in \mathbb{R}^k.$$

Exc: Prove Thm C and Thm D.

Remarks: Suppose X_1, X_2, X_3, X_4 are ind r.v.s. Then the following facts can be shown to hold:

(i) ~~Any~~ The r.v.s in any subcollection ~~subset~~ of $\{X_1, X_2, X_3, X_4\}$ are ind.
In particular $X_i \perp\!\!\!\perp X_j \quad \forall \; 1 \le i < j \le 4$.

(ii) $X_1^2, \; e^{X_2}, \; \log(1+|X_3|), \; \sin X_4$ are ind.

(iii) $X_1 + X_2 + X_3 \perp\!\!\!\perp X_4$.

(iv) $X_1^2 + X_2^2 \perp\!\!\!\perp X_3^7 + e^{X_4}$.

Exc: Prove (i). Also verify (ii) and (iii) in the discrete case.

Exc: Verify (iii) ~~in the~~ when X_1, X_2, X_3 and X_4 are jointly continuous. [Hint: Use (✱)]

The phrase independent and identically distributed (or i.i.d. or iid) has the same meaning. It means that a bunch of r.v.s are independent and they all have the same distribution with same parameter values.

~~Thm:~~ Back to additivity of gamma distn.

Thm: Suppose $X_1, X_2, \ldots, X_k$ are ind r.v.s such that $X_i \sim \text{Gamma}(\alpha_i, \lambda)$ for each $i \in \{1, 2, \ldots, k\}$. Then

$$X_1 + X_2 + \cdots + X_k \sim \text{Gamma}(\alpha_1 + \alpha_2 + \cdots + \alpha_k, \lambda).$$

Proof: Exc [Hint: Use the proposition stated at the end of Pg 93 + induction on k + Remark (iii) of Pg 103.]

Cor1: If $X_1, X_2, \ldots, X_k \overset{iid}{\sim} Exp(\lambda)$, then we have

$$X_1 + X_2 + \cdots + X_k \sim Gamma(k, \lambda).$$

Proof: Apply the theorem stated in Pg 104 in the special case $\alpha_1 = \alpha_2 = \cdots = \alpha_k = 1$.

Cor2: If $Z_1, Z_2, \ldots, Z_k \overset{iid}{\sim} N(0,1)$, then

$$Z_1^2 + Z_2^2 + \cdots + Z_k^2 \sim Gamma\left(\frac{k}{2}, \frac{1}{2}\right).$$

Proof: Exc [**Hint:** See Pg 94 for the $k=2$ case.]

Defn: $Gamma\left(\frac{k}{2}, \frac{1}{2}\right)$ distribution is also called chi-squared distribution (also chi-square ~~or~~ distn. or χ^2-distn) with k degrees of freedom.

Remarks: ① The phrase "k degrees of freedom" is used because there are k independent (and hence "free") standard normal r.v.s in the background (i.e., $Z_1, Z_2, \ldots, Z_k \overset{iid}{\sim} N(0,1)$ as in Cor 2 above) whose sum of squares follow this distn.

② If $X \sim \text{Gamma}\left(\frac{k}{2}, \frac{1}{2}\right)$, then we also use the notation $X \sim \chi^2_k$.

③ In light of Remark ② above, Cor-2 can be restated as

$$Z_1, Z_2, \ldots, Z_k \overset{iid}{\sim} N(0,1) \Rightarrow \sum_{i=1}^{k} Z_i^2 \sim \chi^2_k.$$

④ Chi-squared distⁿ plays a very important role in statistics and related disciplines.

Till now, we were discussing the distⁿ of sum of two (or more) ~~join~~ jointly cont r.v.s. We shall now deal with another scalar valued function of a cont random vector - namely, the ratio.

The distribution of ratio of two jointly cont r.v.s

Suppose (X, Y) is a cont random vector with a joint pdf $f_{X,Y}$. In particular, Y is a cont r.v. and hence it satisfies $P[Y=0]=0$.

Therefore $Z := \frac{X}{Y}$ is a well-defined r.v.

<u>Goal:</u> To find the dist$^{\underline{n}}$ of Z.

<u>Thm:</u> If (X, Y) is a cont random vector with a joint pdf $f_{X,Y}(x,y)$, then $Z := \frac{X}{Y}$ is ~~also~~ a cont r.v. with a pdf

$$f_Z(z) = \int_{-\infty}^{\infty} |y| \, f_{X,Y}(yz, y) \, dy \, , \quad z \in \mathbb{R}.$$

~~If $X \perp\!\!\!\perp Y$~~

<u>Cor:</u> If X, Y are ind cont r.v.s with pdfs f_X, f_Y respectively, then $Z := \frac{X}{Y}$ is also a cont r.v. with a pdf

$$f_Z(z) = \int_{-\infty}^{\infty} |y| \, f_X(yz) \, f_Y(y) \, dy, \quad z \in \mathbb{R}.$$

<u>Cor:</u> In the setup of the above thm, if $\text{Range}(Y) \subseteq (0, \infty)$, and ~~then~~ $\text{Range}(X) \subseteq (0,\infty) = \mathbb{R}^+$,

then $$f_Z(z) = \int_0^{\infty} y \, f_{X,Y}(yz, y) \, dy \, , \quad z \in \mathbb{R}^+.$$

Proof of Thm: We need to show that $\forall\, a \in \mathbb{R}$,

$$(ज)\ldots \quad P(Z \leq a) = \int_{-\infty}^{a} h(z)\, dz \,,$$

where $\quad h(z) = \int_{-\infty}^{\infty} |y|\, f_{X,Y}(yz, y)\, dy \,, \quad z \in \mathbb{R}.$

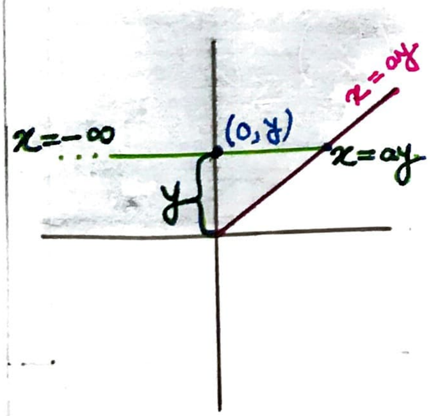
~~Fix $z \in \mathbb{R}$.~~ Take $a \in \mathbb{R}$.

$$\text{LHS of } (ज) = P(Z \leq a) = P\left(\frac{X}{Y} \leq a\right)$$

$$= \overbrace{P(X \leq aY,\ Y > 0)}^{\text{I}} + \overbrace{P(X \geq aY,\ Y < 0)}^{\text{II}}.$$

$$[\because P(Y=0) = 0]$$

The first term $=$ I $= P(X \leq aY,\ Y > 0)$



$$\overset{(*)}{=} \iint_{\substack{y>0 \\ x \leq ay}} f_{X,Y}(x, y)\, dx\, dy$$

$$= \int_{0}^{\infty} \int_{-\infty}^{ay} f_{X,Y}(x, y)\, dx\, dy,$$

Note that in the inner integral, y is a constant.

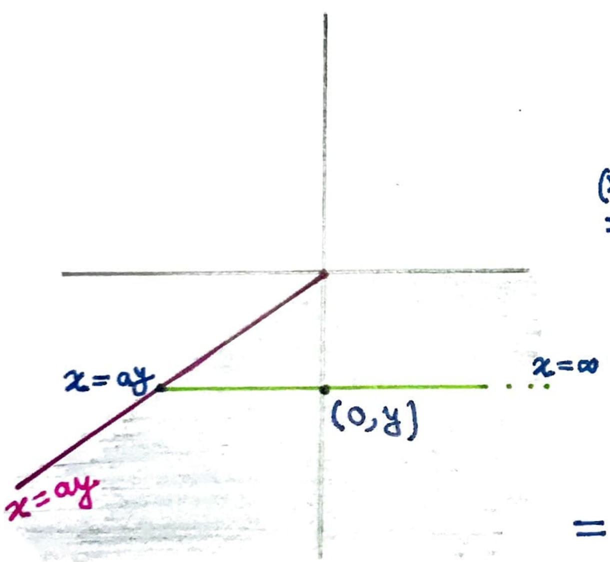
Put $x = yz$ in the inner integral.

$\Rightarrow dx = y\,dz$ and $z = \frac{x}{y}$

$$\Rightarrow \text{I} = \int_0^\infty \int_{-\infty}^{a} y\, f_{X,Y}(yz, y)\, dz\, dy.$$

Similarly, the second term $= \text{II}$

$$= P[X \geq aY,\ Y < 0]$$



$$\overset{(\ast)}{=} \iint_{\substack{y<0 \\ x \geq ay}} f_{X,Y}(x,y)\, dx\, dy$$

$$= \int_{-\infty}^{0} \int_{ay}^{\infty} f_{X,Y}(x,y)\, dx\, dy$$

Again put $x = yz$ in the inner integral.

$\Rightarrow dx = y\,dz$ and $z = \frac{x}{y}$ (but $y<0$).

$$\Rightarrow \text{II} = \int_{-\infty}^{0} \int_{a}^{-\infty} y\, f_{X,Y}(yz, y)\, dz\, dy$$

$$= \int_{-\infty}^{0} \int_{-\infty}^{a} (-y)\, f_{X,Y}(yz, y)\, dz\, dy$$

Therefore, LHS of $(\divideontimes)$

$$= \text{(I)} + \text{(II)}$$

$$= \int_0^\infty \int_{-\infty}^a y f_{X,Y}(yz, y)\, dz\, dy + \int_{-\infty}^0 \int_{-\infty}^a (-y) f_{X,Y}(yz, y)\, dz\, dy$$

$$= \int_0^\infty \int_{-\infty}^a |y| f_{X,Y}(yz, y)\, dz\, dy + \int_{-\infty}^0 \int_{-\infty}^a |y| f_{X,Y}(yz, y)\, dz\, dy$$

$$= \int_{-\infty}^\infty \int_{-\infty}^a |y| f_{X,Y}(yz, y)\, dz\, dy$$

$$\overset{\text{Fubini}}{=} \int_{-\infty}^a \int_{-\infty}^\infty |y| f_{X,Y}(yz, y)\, dy\, dz$$

$$= \int_{-\infty}^a h(z)\, dz$$

$$= \text{RHS of } (\divideontimes).$$

This completes the proof.