Probabilidad y Estadística

Clase 5

Gonzalo Blanco

2023

1. Ley de los grandes números

Esta ley nos dice que cuando el número de repeticiones de un experimento aleatorio ε es muy grande, entonces la frecuencia reltiva de un evento A, dada por

$$f_A = \frac{n_a}{n} = \frac{\textit{cantidad de veces que ocurrió } A}{\textit{cantidad de veces que realicé el experimento}}$$

tiende a la proabilidad real del evento A, P(A). Es decir

$$\lim_{n \to \infty} f_A = P(A)$$

Por lo tanto, si queremos determinar la probabilidad de una evento A de interés, podríamos realizar el experimento aleatori ε una cantidad n de veces y contar la cantidad de veces que ocurrió el evento A, n_a. A partir del cálculo de la frecuencia relativa $f_A = \frac{n_a}{n}$, y según el teorema anterior si hago el experimento muchas veces, es decir si n es muy grande entonces esta frecuencia relativa tiende a la probabilidad real del evento A que llamamos P(A).

Ejercicio para charlar en clase:

Pensar, en función de los conocimientos adquiridos hasta el momento, si existe alguna forma de resolver la situación que se plantea a continuación.

Supongamos que un conocido, amante de los juegos de mesa, sabe que sos estudiantes de probabilidad y estadística; y te pide ayuda para resolver el siguiente problema.

"*Hace poco se sumó al grupo un nuevo integrante. Desde que llegó no para de ganar, creemos que los dados que utiliza están cargados. La última vez que nos juntamos intercambié sus dados por otros para estudiarlos en casa, pero no se como hacer para saber si realmente están cargados o son dados normales. ¿ Se te ocurre alguna forma de ayudarlo ?* "

2. Variable Aleatoria Bidimensional

A veces en un experimento aleatorio nos puede interesar observar dos o más características del mismo. Por ejemplo en un ensayo mecánico sobre un tornillo podriamos necesitas medir presión y temperatura (P,T).

Definición: Sea ε un exp. aleatorio y S su espacio muestral, definimos a $X = X(S)$ e $Y = Y(S)$ dos funciones que asignan un número real a cada uno de los elementos de S.

$\Longrightarrow$ llamamos variable aleatoria bidimensional al par (X, Y).

Si en ves de dos funciones defino n entonces puedo definir tambíen una variable aleatoria n-dimensional. En este curso nos limitaremos al uso de v.a bidimensionales.

Cada una de estas variables aleatorias X e Y tendrán su rango asociado R_x y R_y, y la variable aleatoria bidimensional (X, Y) tendrá su correspondiente rango R_{xy}.

$R_{xy} = \{(x_i, y_j)/x_i \in R_x \ \wedge \ y_j \in R_y \}$

Como cualquier variable aleatoria, la bidimensional también se puede clasificar en discreta o en continua dependiendo de como sean X e Y. Si ambas son discretas entonces (X, Y) será discreta y si ambas son continuas entonces la v.a (X, Y) será continua. No veremos el caso en el cual una v.a es discreta (X) y la otra es continua (Y).

- Si (X, Y) es una v.a discreta entonces $R_{xy} = \{(x_i, y_j) \ \textit{con i=1,2...,n} \wedge \textit{j=1,2,...,m}\}$. Con n y m pertenecientes a los naturales.

- Si (X, Y) es una v.a continua entonces $R_{xy} = \{(x, y)/a < x < b \ \ y \ \ c < y < d\}$. Siendo a,b,c y d números reales.

A partir de ahora daremos tanto las definiciones para el caso de v.a bidimensional discreta como para el caso de continua.

Definición:

Discreta: Sea (X,Y) una v.a discreta. A cada resultado posible (x_i, y_j) le asociamos un número real $p(x_i, y_j)$ que prepresenta la probabilidad de que $X = x_i$ e $Y = y_j$, $P(X = x_i, Y = y_j)$, que satisface las siguientes propiedades:

1. $p(x_i, y_j) \geq 0, \ \forall \, i, j$
2. $\sum\limits_{i=1}^{\infty} \sum\limits_{j=1}^{\infty} p(x_i, y_j) = 1$

Continua: Sea (X,Y) una v.a bidimensional continua, entonces definimos a $f(x, y)$ como la función de densidad de probabilidad conjunta de (X,Y) que satisface:

1. $f(x, y) \geq 0, \ \forall \, (x, y) \in R_{xy}$
2. $\iint\limits_{R_{xy}} f(x, y) dx dy = 1$

Tanto para el caso discreto como para el caso continuo podemos definir la función de probabilidad acumulada o la función de densidad de probabilidad acumulada, respectivamente. Que viene dada por,

$$F(x,y) = P(X \leq x, Y \leq y)$$

Para el caso continuo vale que $\frac{\partial^2 F(x,y)}{dxdy} = f(x,y)$, siempre que F(x,y) sea diferenciable.

2.1. Distribución de probabilidad marginal

Con cada variable aleatoria bidimensional (X, Y) asociamos dos variables aleatorias unidimensionals X e Y. Es decir que también estamos interesados en la función de probabilidad de X y de Y.

$\Longrightarrow$ Definimos lo que se conoce como *Función de probabilidad marginal de X* y *Función de probabilidad marginal de Y*. Que no son más que la función de probabilidad de X y la función de probabilidad de Y.

Discreta:

Función de probabilidad marginal de X $\Longrightarrow p(x_i) = P(X = x_i, -\infty \leq Y \leq \infty) = \sum_{j=1}^{\infty} p(x_i, y_j)$

Función de probabilidad marginal de Y $\Longrightarrow q(y_i) = P(-\infty \leq X \leq \infty, Y = y_j) = \sum_{i=1}^{\infty} p(x_i, y_j)$

Continua:

Función de probabilidad marginal de X $\Longrightarrow f_x(x)$

$P(a \leq X \leq b, -\infty \leq Y \leq \infty) = \int_a^b \left(\int_{-\infty}^{\infty} f(x,y)dy \right) dx = \int_a^b f_x(x)dx$

La función $f_x(x)$ es la función de probabilidad maginal de X.

Función de probabilidad marginal de Y $\Longrightarrow f_y(y)$

$P(-\infty \leq X \leq \infty, c \leq Y \leq d) = \int_c^d \left(\int_{-\infty}^{\infty} f(x,y)dx \right) dy = \int_c^d f_y(y)dy$

La función $f_y(y)$ es la función de probabilidad maginal de y.

Estas funciones de probabilidad marginal, no son más que las funciones de probabilidad o de densidad de probabilidad de cada v.a individual, es decir $f_x(x)$ es la función de densidad de probabilidad de X que llamaremos g(x) y $f_y(y)$ es la función de densidad de probabilidad de Y, que llamaremos h(y). Vale lo mismo para el caso discreto.

Entonces podriamos determinar a partir de f(x,y) las funciones de densidad de probabilidad unidimensionales de X y de Y (vale también para el caso discreto).

2.2. Probabilidad condicional

Discreto:

$P(X = x_i/Y = y_j) = \frac{P(X=x_i, Y=y_j)}{P(Y=y_j)} \quad si \ P(Y = y_j) \neq 0.$

La probabilidad condicional para una v.a bidimensional, cumple que:

$\sum_{i=1}^{\infty} P(x_i/y_j) = \sum_{i=1}^{\infty} \frac{p(x_i,y_j)}{q(y_j)} = \frac{q(y_j)}{q(y_j)} = 1$

Lo mismo para $P(y_j/x_i)$.
$\sum\limits_{j=1}^{\infty} P(y_j/x_i) = 1$

Continuo:
$g(x/y) = \frac{f(x,y)}{h(y)} \quad con\ h(y) > 0$
$h(y/x) = \frac{f(x,y)}{g(x)} \quad con\ g(x) > 0$

La probabilidad condicional para una v.a bidimensional, cumple que:
$\int\limits_{-\infty}^{\infty} g(x/y)dx = \int\limits_{-\infty}^{\infty} \frac{f(x,y)}{h(y)}dx = \frac{1}{h(y)} \int\limits_{-\infty}^{\infty} f(x,y)dx = \frac{h(y)}{h(y)} = 1$
Lo mismo para h(y/x).
$\int\limits_{-\infty}^{\infty} h(y/x)dy = 1$

Ejemplo: Sea(X,Y) una v.a bidimensional cuya función de probabilidad conjunta viene dada por $f(x,y) = x^2 + \frac{xy}{3}$ con $0 \leq x \leq 1$, $0 \leq y \leq 2$.

a) Determinar F(x,y).
$F(x,y) = P(X < x, Y < y) = \int\limits_{-\infty}^{x} \int\limits_{-\infty}^{y} f(t,k)dtdk = \int\limits_{0}^{x} \int\limits_{0}^{y} (t^2 + \frac{tk}{3})dkdt = \int\limits_{0}^{x} (t^2y + \frac{ty^2}{6})dt = \frac{x^3y}{3} + \frac{x^2y^2}{12}$

b) Encontrar la función de probabilidad marginal de X y de Y.
$f_x(x) = g(x) = \int\limits_{0}^{2} (x^2 + \frac{xy}{3})dy = x^2y + \frac{xy^2}{6}\Big|_0^2 = 2x^2 + \frac{2}{3}x$
$f_y(y) = h(y) = \int\limits_{0}^{1} (x^2 + \frac{xy}{3})dx = \frac{x^3}{3} + \frac{yx^2}{6}\Big|_0^1 = \frac{1}{3} + \frac{y}{6}$

c) Calcular g(x/y=0).
$g(x/y=0) = \frac{f(x,0)}{h(0)} = \frac{x^2}{1/3} = \frac{x^2}{3}x$

3. Independencia estadística

X e Y son independientes si el resultado de X no afecta al de Y y viceversa. Entonces vale que,

$$g(x/y) = g(x) \quad \forall y \in R_y$$
$$h(y/x) = h(y) \quad \forall x \in R_x$$

Si X e Y son independientes entonces vale que,

$$f(x,y) = g(x)h(y)$$

.

Lo mismo vale para el caso discreto,

$$p(x,y) = p(x)q(y)$$

Si se conoce f(x,y) podemos determinar g(x) y h(y) pero no al revés. Solo podremos determinar la función de probabilidad conjunta f(x,y) a partir de sus marginales g(x) y h(y) cuando X e Y sean independientes. Lo mismo vale para el caso de v.a discreta. Podremos determinar p(x,y) a partir de p(x) y q(y) si X e Y son independientes.

<u>Ejemplo:</u> Sea (X,Y) una v.a bidimensional con $f(x,y) = e^{-(x+y)}$ con $x \geq 0$ e $y \geq 0$. ¿ Son X e Y independientes?

Si X e Y son independientes entonces f(x,y)=g(x)h(y).

$g(x) = \int\limits_0^\infty e^{-(x+y)}dy = \int\limits_0^\infty e^{-x}e^{-y})dy = e^{-x}\int\limits_0^\infty e^{-y})dy = e^{-x}(-e^{-y})\Big|_0^\infty = e^{-x}$

$h(y) = \int\limits_0^\infty e^{-(x+y)}dx = \int\limits_0^\infty e^{-x}e^{-y})dx = e^{-y}\int\limits_0^\infty e^{-x})dx = e^{-y}(-e^{-x})\Big|_0^\infty = e^{-y}$

$g(x)h(y) = e^{-x}e^{-y} = e^{-(x+y)} = f(x,y) \Longrightarrow$ X e Y son variables aleatorias independientes.

<u>Ejemplo:</u>

Dos hermanos de 9 y 11 años, Juan y Ezequiel, tienen un accidente andando en bicicleta y uno se fractura la muñeca y otro el tobillo. Supongamos que el médico que los atiende, además de tener conocimientos de medicina es un gran estadísta. Entonces combinando ambos campos, haciendo un análisis del accidente y de la contextura física y estilo de vida de ambos hermanos obtiene la siguiente tabla de probabilidades del tiempo requerido para que ambos hermanos tengan una recuperación completa.

Ezequiel / Juan	12	13	14
14	0.01	0.1	0.15
15	0.09	0.3	0.05
16	0.05	0.2	0.05

1) ¿Cuál es la probabilidad de que ambos hermanos completen su recuperación el mismo día?

2) ¿Cuál es la probabilidad de que Juan se recupere antes que Ezequiel?

3) Encontrar las funciones de probabilidad marginales para el tiempo de recuperación de cada hermano.

4) ¿ Puede considerarse al tiempo de recuperación de cada hermano independiente?.

5) Calcular la probabilidad de que Ezequiel se recupere a los 15 días si Juan se recuperó a los 14.

Defino las variables aleatorias que voy a utilizar

J:"tiempo de recuperación de Juan"

E: "tiempo de recuperación de Ezequiel"

Además se que la tabla es la función de probabilidad conjunta de J y E que llamaré f(J,E).

1) Que ambos hermanos se recuperen el mismo día

f(J=14,E=14)=0.15=15 %.

2) Que Juan se recupere antes que Ezequiel
$P(J < E) = 1 - P(J >= E) = 1 - f(14,14) = 1 - 0{,}15 = 0{,}85$.

3) $p(J_i) = \sum_{k=1}^{3} P(J_i/E_k)$ con i=1,2,3
$p(J_1) = P(J = 12) = f(12,14) + f(12,15) + f(12,16) = 0{,}15$
$p(J_2) = P(J = 13) = f(13,14) + f(13,15) + f(13,16) = 0{,}6$
$p(J_3) = P(J = 14) = f(14,14) + f(14,15) + f(14,16) = 0{,}25$

J	12	13	14
p(j)	0.15	0.6	0.25

$p(E_k) = \sum_{i=1}^{3} P(J_i, E_k)$
$p(E_1) = P(E = 14) = f(12,14) + f(13,14) + f(14,14) = 0{,}26$
$p(E_2) = P(E = 15) = f(12,15) + f(13,15) + f(13,15) = 0{,}44$
$p(E_3) = P(E = 16) = f(12,16) + f(13,16) + f(14,16) = 0{,}3$

E	14	15	16
p(e)	0.26	0.44	0.3

4) Independencia
Si son independientes entonces $f(J,E) = p(j) * p(e)$. Veamos si esto es cierto.
f(14,14)=0.15
p(j=14)*p(e=14)=0.25 * 0.26= 0.065
Como $f(14,14) \neq p(j = 14) * p(e = 14)$ entonces no son independientes.

5) $P(E = 15/J = 14) = \frac{f(14,15)}{p(J=14)} = \frac{0{,}05}{0{,}25} = 0{,}2$

4. Función de una variable aleatoria bidimensional

Pueden existir situaciones donde nos interese estudiar una v.a unidimensional que es función de una v.a bidimensional. Un ejemplo de esto podría ser la v.a (X,Y) que representa el largo y el ancho de una pieza y supongamos la v.a $Z = 2X + 2Y$ que repesenta el perímetro de la pieza. Por supuesto que Z es una v.a ya que es función de otra variable aleatoria.

El caso que más nos va a interesar es cuando queremos estudiar situaciones de la forma Z(X,Y)=X+Y.

Sea la v.a (X,Y) con función de probabilidad conjunta f(x,y) y sea Z=X+Y. Veamos algunas características de Z.

$E[Z] = E[X+Y] = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} Z(x,y)f(x,y)dxdy = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty}(x+y)f(x,y)dxdy = \int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} xf(x,y)dxdy +$
$\int\limits_{-\infty}^{\infty}\int\limits_{-\infty}^{\infty} yf(x,y)dxdy = \int\limits_{-\infty}^{\infty} x\left[\int\limits_{-\infty}^{\infty} f(x,y)dy\right]dx + \int\limits_{-\infty}^{\infty} y\left[\int\limits_{-\infty}^{\infty} f(x,y)dx\right]dy = \int\limits_{-\infty}^{\infty} xg(x)dx +$
$\int\limits_{-\infty}^{\infty} yh(y)dy = E[X] + E[Y]$

$\Longrightarrow E[X+Y] = E[X] + E[Y]$

Esto se puede generalizar para la suma de n variables aleatorias X_i.

$$E[X_1 + X_2 + X_3 + ... + X_n] = E[X_1] + E[X_2] + E[X_3] + ... + E[X_n]$$

Veamos ahora cual es la varianza de Z.

$V(Z) = E[Z^2] - (E[Z])^2 = E[(X+Y)^2] - (E[X+Y])^2 = E[X^2 + Y^2 + 2XY] - (E[X] + E[Y])^2 =$
$= E[X^2] + E[Y^2] + 2E[XY] - E[X]^2 - E[Y]^2 - 2E[X]E[Y] =$
$= E[X^2] - E[X]^2 + E[Y^2] - E[Y]^2 + 2\{E[XY] - E[X]E[Y]\} = V(X) + V(Y) + 2Cov(X,Y)$

$\Longrightarrow V(X+Y) = V(X) + V(Y) + 2Cov(X,Y)$

Siendo $Cov(X,Y) = E[XY] - E[X]E[Y]$, y se lo conoce como covarianza de (X,Y), y la podemos denotar como σ_{xy}.
La covarianza mide la dependencia estadística entre dos variables aleatorias.
Si X e Y son independientes entonces la covarianza entre X e Y es igual a 0, $Cov(X,Y) = \sigma_{xy} = 0$.

Otra forma de escribir la covarianza entre X e Y es $Cov(X,Y) = E[(X - \mu x)(Y - \mu_y)]$

$Cov(X,Y) = E[(X - \mu x)(Y - \mu_y)] = E[XY - YE[X] - XE[Y] + E[X]E[Y]] = E[XY] - E[Y\mu_x] - E[X\mu_y] - E[\mu_x\mu_y] = E[XY] - \mu_x E[Y] - \mu_y E[X] + \mu_x\mu_y = E[XY] - \mu_x\mu_y - \mu_x\mu_y + \mu_x\mu_y = E[XY] - \mu_x\mu_y = E[XY] - E[X]E[Y]$.

Dijimos entonces que la covarianza nos daba información de la dependencia entre dos variables aleatorias X e Y, ahora bien, podemos definir el grado de dependencia lineal entre dos variables aleatorias a partir de la estimación del **coeficiente de correlación lineal** ρ_{xy}, el cual viene dado por

$$\rho_{xy} = \frac{Cov(X,Y)}{\sigma_x\sigma_y}$$

$$-1 \leq \rho_{xy} \leq 1$$

Este coeficiente solo puede tomar valore sentre [-1,1] y nos dice el grado de dependencia lineal entre X e Y.

Si $\rho_{xy} = 1$ entonces la relación entre X e Y es perfectamente lineal, es decir que podemos escribir a Y como Y=aX+b, con $a > 0$

Si $\rho_{xy} = -1$ entonces la relación entre X e Y es perfectamente lineal inversa, es decir que podemos escribir a Y como Y=-aX+b, con $a > 0$.

Si X e Y son independientes entonces $\rho_{xy} = 0$.

Observación: No vale la inversa, que $\rho_{xy} = 0$ no implica necesariamente que X e Y sean independientes, solo que su dependencia no es lineal.

Veamos las siguientes figuras donde se grafican dos variables aleatorias X e Y.

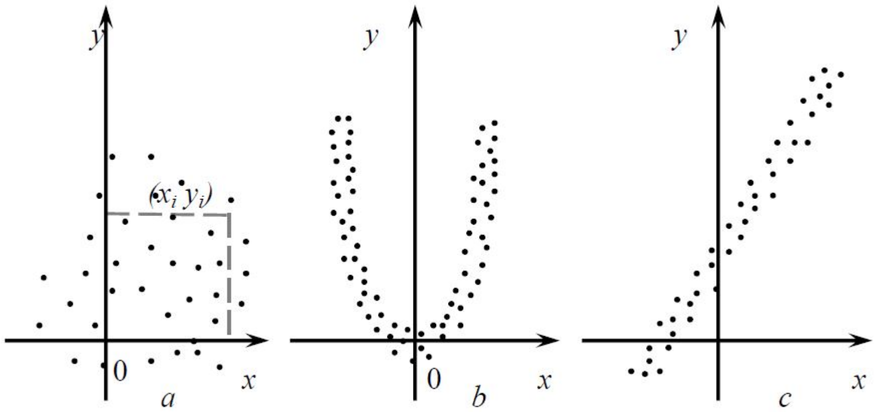

Figura 1: Gráficos de X vs Y

En la figura a) podemos ver que no existe ningún tipo de dependencia entre X e Y, no se observa alguna relación funcional, su coeficiente de correlación lineal será cercano a 0.

En la figure b) se observa una dependencia entre X e Y, y si miramos con atención paracería del tipo $Y = X^2$, pero no la podremos determinar con ρ_{xy} ya que este coeficiente solo nos da información del grado de dependencia lineal.

El caso c) es el que nos interesa caracterizar, se observa claramente una dependencia lineal entre X e Y, si calculamos el coef. de correlación lineal dará muy cercano a 1.

Ahora bien, que ρ_{xy} sea muy chico o cercano a 0 no implica que no exista una dependencia entre X e Y, puede que exista una relación funcional pero que no sea lineal como el caso b).

Resumamos entonces las propiedades que fuimos nombrando y agregemos otras,

Propiedades: Sean X,Y y Z v.a y a,b,c y d constantes reales.

- Cov(X,X)=V(X)
- Cov(X+Y,Z)=Cov(X,Z)+Cov(Y,Z)
- Cov(a+bX,c+dY)= b d Cov(X,Y)
- Si X e Y son independientes $\Longrightarrow \rho_{xy} = 0$, no necesariamente se cumple la inversa.
- $-1 \leq \rho_{xy} \leq 1$

Demostración:

Sea la v.a $\frac{x}{\sigma_x} + \frac{y}{\sigma_y}$

$0 \leq V(\frac{x}{\sigma_x} + \frac{y}{\sigma_y}) = \frac{1}{\sigma_x^2}V(X) + \frac{1}{\sigma_y^2}V(Y) + 2\frac{Cov(X,Y)}{\sigma_x\sigma_y} = 1 + 1 + 2\rho_{xy} = 2(1 + \rho_{xy})$

$0 \leq 2(1 + \rho_{xy}) \Longrightarrow \rho_{xy} \geq -1$

Defino ahora la v.a $\frac{x}{\sigma_x} - \frac{y}{\sigma_y}$
$0 \leq V(\frac{x}{\sigma_x} - \frac{y}{\sigma_y}) = 1 + 1 - 2\rho_{xy} = 2(1 - \rho_{xy})$
$0 \leq 2(1 - \rho_{xy}) \Longrightarrow \rho_{xy} \leq 1$

$\Longrightarrow -1 \leq \rho_{xy} \leq 1$

• Si $\rho_{xy} = 1$ entonces Y=aX+b, con $a > 0$.
• Si $\rho_{xy} = -1$ entonces Y=-aX+b, con $a > 0$.

Se define la **matriz de covarianza** de X,Y como
$C_{xy} = \begin{pmatrix} Cov(X,X) & Cov(X,Y) \\ Cov(X,Y) & Cov(Y,Y) \end{pmatrix} = \begin{pmatrix} \sigma_x^2 & \sigma_{xy} \\ \sigma_{xy} & \sigma_y^2 \end{pmatrix} = \begin{pmatrix} \sigma_x^2 & \rho_{xy}\sigma_x\sigma_y \\ \rho_{xy}\sigma_x\sigma_y & \sigma_y^2 \end{pmatrix}$

y se la conoce como matriz de errores. Esta matriz contiene los datos de dispersión de cada v.a así como la dependencia entre cada una de estas variables aleatorias. Se utilizará para calcular por ejemplo la dispersión de una v.a Z que es de la forma Z(X,Y).

Eso se puede generalizar para n v.a X_i, entonces la matriz de covarianza C de todas estas v.a quedaria de la forma,

$$C = \begin{pmatrix} \sigma_1^2 & \rho_{1,2}\sigma_1\sigma_2 & ... & \rho_{1,n}\sigma_1\sigma_n \\ \rho_{2,1}\sigma_2\sigma_1 & \sigma_2^2 & ... & \rho_{2,n}\sigma_2\sigma_n \\ ... & ... & ... & ... \\ \rho_{n,1}\sigma_n\sigma_1 & ... & ... & \sigma_n^2 \end{pmatrix}$$

5. Ley de Propagación de errores

Supongamos que tenemos n variables aleatorias $(X_1, X_2, ..., X_n) = \bar{X}$, y una variable aleatoria $Y = Y(X_1, X_2, ..., X_n)$. Asumimos que también conocemos los valores medios de cada X_i dados por μ_i y la matriz de covarianza C. Quiero determinar V(Y).

Para obtenerlo partiremos de un desarrollo de taylor de $Y(\bar{X})$ alrededor de $\bar{\mu} = (\mu_1, \mu_2, ..., \mu_n)$

$Y(\bar{X}) = Y(\bar{\mu}) + \sum_{i=1}^{n}(X_i - \mu_i)\frac{\partial Y}{\partial X_i}\Big|_{\bar{X}=\bar{\mu}} +$ términos de orden superior
Si tomo $E[Y(\bar{X})]$
$E[Y(\bar{X})] = E[Y(\bar{\mu}) + \sum_{i=1}^{n}(X_i - \mu_i)\frac{\partial Y}{\partial X_i}\Big|_{\bar{X}=\bar{\mu}} +$ términos de orden superior]
$E[Y(\bar{X})] = E[Y(\bar{\mu})] + \frac{\partial Y}{\partial X_i}\Big|_{\bar{X}=\bar{\mu}} \sum_{i=1}^{n} E[(X_i - \mu_i)] +$ E[términos de orden superior]
$E[Y(\bar{X})] = Y(\bar{\mu}) +$ E[términos de orden superior]

Asumiendo que $X_i - \mu_i$ es pequeño $\Longrightarrow$ los términos de orden superior son despreciables respecto a $Y(\bar{\mu})$.

$\Longrightarrow E[Y(\bar{X})] = Y(\bar{\mu})$

Calculemos ahora $V(Y(\bar{X}))$.

$V(Y(\bar{X})) = E[(Y(\bar{X}) - E[Y(\bar{X})])^2] = E[(Y(\bar{X}) - Y(\bar{\mu}))^2]$
Del desarrollo anterior podemos obtener,
$Y(\bar{X}) - Y(\bar{\mu})) = \sum_{i=1}^{n}(X_i - \mu_i)\frac{\partial Y}{\partial X_i}\Big|_{\bar{X}=\bar{\mu}}$
Entonces,
$V(Y(X)) = E[\left(\sum_{i=1}^{n}(X_i - \mu_i)\frac{\partial Y}{\partial X_i}\Big|_{\bar{X}=\bar{\mu}}\right)^2] =$
$= E[\sum_{i=1}^{n}\sum_{j=1}^{n}(X_i - \mu_i)(X_j - \mu_j)\frac{\partial Y}{\partial X_i}\Big|_{\bar{X}=\bar{\mu}}\frac{\partial Y}{\partial X_j}\Big|_{\bar{X}=\bar{\mu}}] =$
$V(Y(\bar{X})) = \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial Y}{\partial X_i}\Big|_{\bar{X}=\bar{\mu}}\frac{\partial Y}{\partial X_j}\Big|_{\bar{X}=\bar{\mu}}E[(X_i - \mu_i)(X_j - \mu_j)]$
$V(Y(\bar{X})) = \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial Y}{\partial X_i}\Big|_{\bar{X}=\bar{\mu}}\frac{\partial Y}{\partial X_j}\Big|_{\bar{X}=\bar{\mu}}Cov(X_i, X_j)$

$$V(Y(\bar{X})) = \sum_{i=1}^{n}\sum_{j=1}^{n}\frac{\partial Y}{\partial X_i}\Big|_{\bar{X}=\bar{\mu}}\frac{\partial Y}{\partial X_j}\Big|_{\bar{X}=\bar{\mu}}C_{ij}(\bar{X})$$

Esta fórmula es lo que se conoce como Ley de Propagación de errores y es válida para valores de $\bar{X}$ cercanos a $\bar{\mu}$ y siempre que $\sigma_{x_i} < \mu_{x_i}$, ya que si el desvió estándar de cada v.a es del orden de su media entonces no puedo despreciar los términos de orden superior en el desarrollo de Taylor y esta aproximación deja de ser válida.

Si mis X_i son v.a independientes entre si $\Longrightarrow C_{ij} = 0 \quad \forall i \neq j$.

$$\Longrightarrow V(Y(\bar{X})) = \sum_{i=1}^{n}\left(\frac{\partial Y}{\partial x_i}\Big|_{\bar{X}=\bar{\mu}}\right)^2 C_{ii}^2$$

$$V(Y(\bar{X})) = \sum_{i=1}^{n}\left(\frac{\partial Y}{\partial X_i}\Big|_{\bar{X}=\bar{\mu}}\right)^2 \sigma_i^2$$

<u>Ejemplo:</u> Supongamos un conjunto de n v.a $X_1, X_2, X_3, ..., X_n$. Todas con la misma varianza σ^2 y la misma media μ.

Sea $Y = \frac{1}{n}\sum_{i=1}^{n}X_i = \bar{X}$

$E[Y] = \frac{1}{n}\sum_{i=1}^{n}E[X_i] = \frac{1}{n}n\mu = \mu$, es decir que Y tiene el mismo valor esperado que cada una de mis X_i.

Aplicando la ley de prop. de errores llego a,
$V(Y) = \sum_{i=1}^{n}\left(\frac{\partial Y}{\partial X_i}\right)^2\sigma^2 = \frac{\sigma^2}{n}$

Este resultado es muy útil, veamos la razón. Supongamos que en un experimento mido alguna variable de interés (presión, temperatura, etc) con un sensor que tiene una precisión o un error de medición que podemos asociarlo al desvío estándar y denotarlo σ.

Supongamos entonces que mido 10 veces esta variable de interés y obtengo $(P_1, P_2, P_3, ..., P_{10})$. Ahora bien, cada una de estas mediciones tiene el mismo error σ asociado, ya que se realizaron con el mismo sensor.

Si defino la v.a $P_m = \sum_{i=1}^{10} \frac{P_i}{10}$

$\Longrightarrow V(P_m) = \sigma_m^2 = \frac{\sigma^2}{10}$

$\Longrightarrow \sigma_m = \sqrt{\frac{\sigma_m^2}{10}} = \frac{\sigma_m}{\sqrt{10}}$

Es decir que el promedio simple de mis 10 observaciones disminuyó el error de mi variable de interés.

Entonces si tengo un sensor con cierto error de medición, es conveniente tomar como dato final a un promedio de n cantidad de datos ya que mientras mayor datos tengo, más disminuye mi error de medición. Si $n \longrightarrow \infty$ entonces $\sigma \longrightarrow 0$.