§4.7:

Normal Approximation to the Binomial and Poisson Distributions:

If X is binomial random variable

(تذكر) $\mu = np$ $\sigma^2 = np(1-p)$

$\rightarrow$ approximate to normal $\Rightarrow Z = \dfrac{X-\mu}{\sigma} = \dfrac{X-np}{\sqrt{np(1-p)}}$

* Contininuity correction: (± 0.5)

$P(X \leq x) = P(X \leq x + 0.5)$

$P(X \geq x) = P(X \geq x - 0.5)$

① Example: Assume that in a digital communication channel, the number of bits received in error can be modeled by a binomial random variable, and assume that the probability that a bit is ~~rese~~ received in error is 10^{-5} (p). If 16 million (n) bits are transmitted, what is the probability that 150 or fewer errors occure?

$P = 10^{-5}$ $n = 16\text{million} = 16 \times 10^6$ $P(X \leq 150) = ??$

1- By binomial:

$P(X \leq 150) = \sum_{x=0}^{150} \binom{16\times10^6}{x} (10^{-5})^x (1-10^{-5})^{16\times10^6 - x} \rightarrow$ استخرج x لأنه cumulative

بنجمع من $x=0 \rightarrow x=150$ ⇐ مستحيل ⇐ بنروح على approximation

2- By Normal Approximation:

$P(X \leq 150) = P(X \leq 150 + 0.5) = P(X \leq 150.5)$

$P\left(Z \leq \dfrac{150.5 - (16\times10^6)(10^{-5})}{\sqrt{(16\times10^6)(10^{-5})(1-10^{-5})}}\right) = P\left(Z \leq \dfrac{150.5 - 160}{12.649}\right)$

$P(Z \leq -0.751) = 0.227 \longrightarrow$ from table

* normal approximation is work at n large relative to p. as the previouse example

* binomial approximation is affective if $\dfrac{n}{N} < 0.1$ → n = sample size, → N = population size

$np > 5$, $n(1-p) > 5$

Normal approximation for Poisson random variable:

$$Z = \frac{X-\lambda}{\sqrt{\lambda}} \Rightarrow \lambda = E(x) = V(x)$$

* using the same correction used for binomial (± 0.5)

** The approximation is good for $\lambda > 5$

② **Example:** Assume that the number of asbestos particles in a sequared meter of dust on a surface follows a Poisson distribution with a mean λ of 1000. If a sequared meter of dust is analyzed, what is the probability that 950 or fewer particles are found? $P(X \leq 950)$

$$P(X \leq 950) = P(X \leq 950 + 0.5) = P(X \leq 950.5)$$

$$= P\left(Z \leq \frac{950.5 - 1000}{\sqrt{1000}}\right) = P(Z \leq -1.57) = 0.058208$$

§ 4.8: Exponential Distribution:-

Exponential random variable: X = is equals the distance between successive events from a Poisson process with mean number of events $\lambda > 0$ per unit interval, with parameter λ.

The probability density function of X is:

$$f(x) = \lambda e^{-\lambda x} \qquad 0 \leq x < \infty$$

*** Exponential distri. يكون مصاحب لـ Poisson process بس الفرق إنه

Exponential: المسافة بين كل نجاح والتاني في الأحداث.

$$E(x) = \mu = \frac{1}{\lambda} \text{ (mean)} \qquad V(x) = \sigma^2 = \frac{1}{\lambda^2}$$

③

process (وش) distribution → Exponential

Example: In a large corporate computer network, user log-ons to the system can be modeled as a Poisson process with a mean λ of 25 log-ons per hour. ① What is the probability that there are no log-ons in an interval of six minutes?

no log-ons ⇒ failar (failur)

(X) 6 minutes ⇒ $X > 6$ minutes ⇒ $X > 0.1$ hour → حولها للساعة ١

$$P(0.1 < x) = \int_{0.1}^{\infty} \lambda e^{-\lambda x} dx = \int_{0.1}^{\infty} 25 e^{-25x} dx$$

$$= \frac{25}{-25} e^{-25x} \Big]_{0.1}^{\infty} = -e^{-25x} \Big]_{0.1}^{\infty} = e^{-25x} \Big]_{\infty}^{0.1} = e^{-2.5} - e^{-\infty} = e^{-2.5} - 0$$

قلبنا حدود التكامل

$$\boxed{P(0.1 < x) = 0.0821}$$

② * What is the probability that the ~~next~~ time until the next log-ons is between 2 & 3 minutes?

$$P(2min < x < 3min) = P(0.033 < x < 0.05)$$

$$= \int_{0.033}^{0.05} \lambda e^{-\lambda x} dx = \int_{0.033}^{0.05} 25 e^{-25x} dx$$

$$e^{-25x} \Big]_{0.05}^{0.033} = e^{-25 * 0.033} - e^{-25 * 0.05} = 0.438 - 0.287$$

$$= 0.1517$$

③ * Determine the interval of time such that the probability that no log-on occures in the interval is 0.9

time = x in minutes.

$$P(X > x) = e^{-25x} = 0.9$$

$$-25x = \ln 0.9$$

$$x = 4.214 \times 10^{-3} \text{ (hours)} \left| \frac{60 \text{ min}}{1 \text{ hour}} \right.$$

$$x = 0.25 \text{ minute}$$

④ * The mean time until the next log-on & its standard deviation:

$$\mu = \sigma = \frac{1}{\lambda} = \frac{1}{25} = 0.04 \text{ hour} = 2.4 \text{ minutes}$$

Ch#5: Joint Probability Distributions:

§5.1: Two or More Random variables:

joint probability distribution:

The probability distribution that defines X & Y (random variables) & their simultaneous behavior

bivariate probability distribution: The joint probability distribution of 2 random variables, is usually written as $P(X=x, Y=y)$
discrete

joint probability mass function $f_{XY}(x,y)$ satisfies:

(1) $f_{XY}(x,y) \geq 0$

(2) $\sum_x \sum_y f_{XY}(x,y) = 1$

(3) $f_{XY}(x,y) = P(X=x, Y=y)$

* probability outside the range $(X, Y) = 0$

" " range $X = 0$

joint probability density function: (volume under the surface $f_{XY}(x,y)$) statisfies the following properties:

(1) $f_{XY}(x,y) \geq 0$ for all x, y

(2) $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x,y)\,dx\,dy = 1$

(3) $P((X,Y) \in R) = \iint_R f_{XY}(x,y)\,dx\,dy$

R = region of 2D space

4

Example: X: the time until a computer server connects to the machine

Y: the time until the surver authorizes you as a valid user.

Assume that the joint probability density function for X & Y:

$$f_{XY}(x,y) = 6\times10^{-6}\exp(-0.001x - 0.002y) \qquad x<y$$

① show that $f_{XY}(x,y)$ is density function.

⇐ المطلوب: إثبات $f_{XY}(x,y)$ ال density function من خلال اثبات انه $\int_{\infty}\int f = 1$

* في الاقتران حكالك إنه $X<y$ ⇐ أقل قيمة $y = x$

وأعلى قيمة $y = \infty$

بما إنه عدد مرات احسب بالخادم للكمبيوتر عدد لا يمكن أن يكون سالب ⇐ اقل قيمة $x = 0$

أعلى قيمة $x = \infty$

وبما أن y بدلالة x نبدأ التكامل بـ (dy)

$$\int_0^{\infty}\int_x^{\infty} 6\times10^{-6}\, \exp(-0.001x - 0.002y)\, dy\, dx$$

$$6\times10^{-6}\int_0^{\infty}\int_x^{\infty} \exp(-0.001x)\cdot \boxed{\exp(-0.002y)\, dy}\, dx$$

$$\frac{6\times10^{-6}}{-0.002}\int_0^{\infty} \exp(-0.001x)\left[\exp(-0.002*\infty) - \exp(-0.002x)\right] dx$$

$$3\times10^{-3}\int_0^{\infty} \exp(-0.001x)\cdot \exp(-0.002x)\, dx$$

$$3\times10^{-3}\int_0^{\infty} \exp(-0.003x)\, dx = \frac{\cancel{3\times10^{-3}}^{\,-1}}{-\cancel{0.003}}\left[\exp(-0.003*\infty) - \exp(-0.003*0)\right]$$

$= 1$

$$= -1 * -1 = \boxed{1}$$

② The probability that $X < 1000$ & $Y < 2000$?

$X \leftarrow$ أكبر قيمة $x = 1000$ وأقل قيمة $x = 0$

$Y \leftarrow$ أكبر قيمة $y = 2000$ وأقل قيمة $y = x$

$$6\times10^{-6}\int_0^{1000}\int_x^{2000} \exp(-0.001x)\, \exp(-0.002y)\, dy$$

$$\frac{6\times10^{-6}}{-0.002}\int_0^{1000} \exp(-0.001x)\left[\exp(-0.002*2000) - \exp(-0.002x)\right] dx$$

$$3\times10^{-3}\int_0^{1000} \exp(-4)\, \exp(-0.001x) - \exp(-0.003x)\, dx$$

$$3\times10^{-3}\left[\frac{\exp(-4)}{-0.001}\left[\exp(-0.001*1000) - \exp(0)\right] - \frac{\left[\exp(-0.003*1000) - \exp(0)\right]}{-0.003}\right]$$

$$3\times10^{-3}\left[305.16\right] = 0.915$$

Marginal Probability Distributions:

The individual probability distribution of a random variable that can be determined from the joint probability distribution

marginal probability mass function:

$$f_X(x) = \sum_{X=x} \sum_{Y=y_1}^{y_n} f_{XY}(x,y)$$

capital small

⑤ Examples:

y = Response time	x = Number of Bars of Single Strength			Marginal of Y
	1	2	3	
4	0.15	0.1	0.05	0.3
3	0.02	0.1	0.05	0.17
2	0.02	0.03	0.2	0.25
1	0.01	0.02	0.25	0.28
Marginal of X	0.2	0.25	0.55	

$$f_X(3) = P(X=3) = \sum_{X=3} \sum_{y=1}^{4} f_{XY}(x,y)$$

$$= P(X=3, Y=1) + P(X=3, Y=2) + P(X=3, Y=3) + P(X=3, Y=4)$$

$$= 0.25 + 0.2 + 0.05 + 0.05$$

$$= 0.55$$

Marginal probability density function:

$$f_X(x) = \int f_{XY}(x,y)\, dy \qquad f_Y(y) = \int f_{XY}(x,y)\, dx$$

$$P(a < x < b) = \int_a^b \int_{-\infty}^{\infty} f_{XY}(x,y)\, dy\, dx$$

Conditional Probability Distribution:

$$P(Y=B \mid X=A) = \frac{P(X=A, B=Y)}{P(X=A)} \quad \{\text{Mass function}\}$$

Example: (5 continue)

(1) $P(Y=1 \mid X=3) = \frac{P(X=3, Y=1)}{P(X=3)} = \frac{0.25}{0.55} = 0.454$

تقاطع

$P(X=3)$ → Marginal

(2) The probability that $Y=2$ given that $X=3$:

$$P(Y=2 \mid X=3) = \frac{P(X=3, Y=2)}{P(X=3)} = \frac{0.2}{0.55} = \frac{f_{XY}(3,2)}{f_X(3)} = 0.364$$

تقاطع

$f_X(3)$ → Marginal

(3) $P(Y=1 \mid X=3) + P(Y=2 \mid X=3) + P(Y=3 \mid X=3) + P(Y=4 \mid X=3)$

$$= \frac{0.25}{0.55} + \frac{0.2}{0.55} + \frac{0.05}{0.55} + \frac{0.05}{0.55} = \frac{0.55}{0.55} = \boxed{1}$$

Conditional Probability density function:

$$f_{Y|X}(y) = \frac{f_{XY}(x,y)}{f_X(x)} \quad \text{for } f_X(x) > 0$$

Properties:

(1) $f_{Y|X}(y) \geq 0$

(2) $\int f_{Y|X}(y)\, dy = 1$

(3) $P(Y \in B \mid X=x) = \int_B f_{Y|X}(y)\, dy$ $\quad B \to$ range of Y

Example (4) continue:

$$f_{XY}(x,y) = 6\times10^{-6}\exp(-0.001x - 0.002y)$$

(1) Marginal density function of x for $x>0$

لما يُنه طالب Marginal for (x) ⇐ بكامل بالنسبة لـ (y)

حدود التكامل : $y>x$ ⇐ أول قيمة $y=x$

أكبر قيمة $y=\infty$

$$f_X(x) = \int_x^\infty 6\times10^{-6}\exp(-0.001x)\exp(-0.002y)\,dy$$

$$6\times10^{-6}\exp(-0.001x)\int_x^\infty \exp(-0.002y)\,dy$$

$$\frac{6\times10^{-6}}{-0.002}\exp(-0.001x)\left[\overset{0}{\cancel{\exp(-0.002\times\infty)}} - \exp(-0.002x)\right]$$

$$\boxed{f_X(x) = 3\times10^{-3}\exp(-0.003x)}$$

(2) conditional probability density function:

$$f_{Y|x}(y) = \frac{f_{XY}(x,y)}{f_X(x)} = \frac{6\times10^{-6}\exp(-0.001x)\exp(-0.002y)}{3\times10^{-3}\exp(-0.003x)}$$

$$= 2\times10^{-3}\exp(-0.002y)\exp(0.002x)$$

(3) The conditional probability density function of ~~y~~ Y, given that X = 1500, is non zero. Determine the

probability that Y exceeds 2000, given that $x = 1500$

$$P(Y > 2000 \mid x = 1500) = \int_{2000}^{\infty} f_{Y|1500}(y)\, dy$$

$$= \int_{2000}^{\infty} 2\times10^{-3}\ \exp(0.002 \overset{(1500)}{\cancel{x}})\ \exp(-0.002y)\, dy$$

$$2\times10^{-3} \exp(0.002 \overset{(1500)}{\cancel{x}}) \int_{2000}^{\infty} \exp(-0.002y)\, dy$$

$$\frac{2\times10^{-3}}{-0.002}\ \exp(0.002 \times 1500)\ [\exp(-0.002\, \infty) - \exp(-0.002 \times 2000)]$$

(−1) 0

$$-\exp(3) * -\exp(-4) = 0.368$$

Conditional mean & variance:

Conditional mean: $E(Y|x) = \mu_{Y|x} = \int_{y} y\, f_{Y|x}(y)$

conditional variance: $V(Y|x) = \sigma^2_{Y|x} = \int_{y} (y - \mu_{Y|x})^2 f_{Y|x}(y)$

Independence:

if X & Y are independent:

(1) $f_{XY}(xy) = f_X(x) f_Y(y)$ → for all x & y

(2) $f_{Y|x}(y) = f_Y(y)$ → for all x & y with $f_X(x) > 0$

(3) $f_{X|y}(x) = f_X(x)$ → for all x & y with $f_Y(y) > 0$

(4) $P(X \in A; Y \in B) = P(X \in A)\, P(Y \in B)$ → for any A & B in the range X, Y

* إذا وحدة من الخصائص السابقة صحيحة ← كلهم صحيحين

example:

$$f_{XY}(x,y) = 2\times10^{-6} \exp(-0.002y) \exp(-0.001x)$$

$$x \geq 0 \quad \& \quad y \geq 0$$

(1) Show that X & Y are independent.

$$f_{XY}(x,y) = f_X(x)\, f_Y(y)$$

$$f_X(x) = \int_0^\infty \left[2\times10^{-6} \exp(-0.002y) \exp(-0.001x)\right] dy$$

$$= \frac{2\times10^{-6}}{0.002} \exp(-0.001x) \left[\exp(0) - \cancel{\exp(-\infty)}\right] = \boxed{10^{-3} \exp(-0.001x)}$$

$$f_Y(y) = \int_0^\infty 2\times10^{-6} \exp(-0.002y) \exp(-0.001x)\, dx$$

$$= 2\times10^{-3} \exp(-0.002y) \left[\exp(-0.001\times0) - \cancel{\exp(-0.001\times\infty)}\right]$$

$$\boxed{= 2\times10^{-3} \exp(-0.002y)}$$

$$\Rightarrow 2\times10^{-3} \exp(-0.002y) \times 10^{-3} \exp(-0.001x)$$

$$= 2\times10^{-6} \exp(-0.002y) \exp(-0.001x) = f_{XY}(x,y)$$

(2) Determine $P(X>1000, Y\leq 1000)$

$$P(X>1000, Y<1000) = P(X>1000)\, P(Y<1000)$$

$$= \int_{1000}^{\infty} f_X(x)\,dx * \int_{0}^{1000} f_Y(y)\,dy$$

$$= \int_{1000}^{\infty} 10^{-3} \exp(-0.001x)\,dx * \int_{0}^{1000} 2\times10^{-3} \exp(-0.002y)\,dy$$

$$= \frac{10^{-3}}{0.001}\left[\exp(-1) - \exp(-\infty)\right] * \frac{2\times10^{-3}}{0.002}\left(\exp(0) - \exp(-2)\right)$$

$$\exp(-1) * (1 - \exp(-2)) = 0.318$$

Example: Let the random variables X & Y denote the lengths of two dimensions of a machined part, respectively. Assume that X & Y are independent random variables, and further assume that the distribution of X is normal with mean μ 10.5 mm & variance σ^2 0.0025 mm^2 and that distribution of Y is normal with mean μ 3.2 mm & variance σ^2 0.0036

Determine the probability that $\boxed{10.4 < X < 10.6 \text{ \& } 3.15 < Y < 3.25}$??

$$P(10.4 < X < 10.6,\ 3.15 < Y < 3.25) = ??$$

$$= P(10.4 < X < 10.6) * P(3.15 < Y < 3.25)$$

$$= P\left(\frac{10.4-10.5}{\sqrt{0.0025}} < Z < \frac{10.6-10.5}{\sqrt{0.0025}}\right) * P\left(\frac{3.15-3.2}{\sqrt{0.0036}} < Z < \frac{3.25-3.2}{\sqrt{0.0036}}\right)$$

$$P(-2 < Z < 2) * P(-0.83 < Z < 0.83)$$

$$= \{P(Z<2) - P(Z<-2)\} * \{P(Z<0.83) - P(Z<-0.83)\}$$

$$= 0.9542 * 0.5937 = 0.567$$

§ 5.2: Covariance & Correlation:

Expected Values:

$$E[h(X,Y)] = \begin{cases} \sum\sum h(x,y)\, f_{XY}(x,y) \rightarrow X, Y \text{ Discrete} \\ \iint h(x,y)\, f_{XY}(x,y)\,dx\,dy \rightarrow X, Y \text{ continuous} \end{cases}$$

Example: (5 continue)

For the joint probability distribution of the 2 random variables calculate $E[(X-\mu_X)(Y-\mu_Y)]$

$\mu_X = 2.35$ $\mu_Y = 2.49$

X	Y	$X-\mu_X$	$Y-\mu_Y$	$(1-\mu_X)(Y-\mu_Y)$	$(2-\mu_X)(Y-\mu_Y)$	$(3-\mu_X)(Y-\mu_Y)$
1	1	-1.35	-1.49	2.0115	0.5215	-0.9685
2	2	-0.35	-0.49	0.6615	0.1715	-0.3185
3	3	0.65	0.51	-0.6885	-0.1785	0.3315
	4		1.51	-2.0385	-0.5285	0.9815

$$\sum\sum (x-\mu_X)(Y-\mu_Y)\, f_{XY}(x,y) = 2.0115*0.01 + 0.6615*0.02$$
$$+ -0.6885*0.02 + -2.0385*0.15 + 0.5215*0.02 + 0.6615*0.03$$
$$+ -0.1785*0.1 - 0.5285*0.1 - 0.9685*0.25 - 0.3185*0.2 + 0.3315*0.05$$
$$+ 0.05*0.9815$$
$$= -0.5815$$

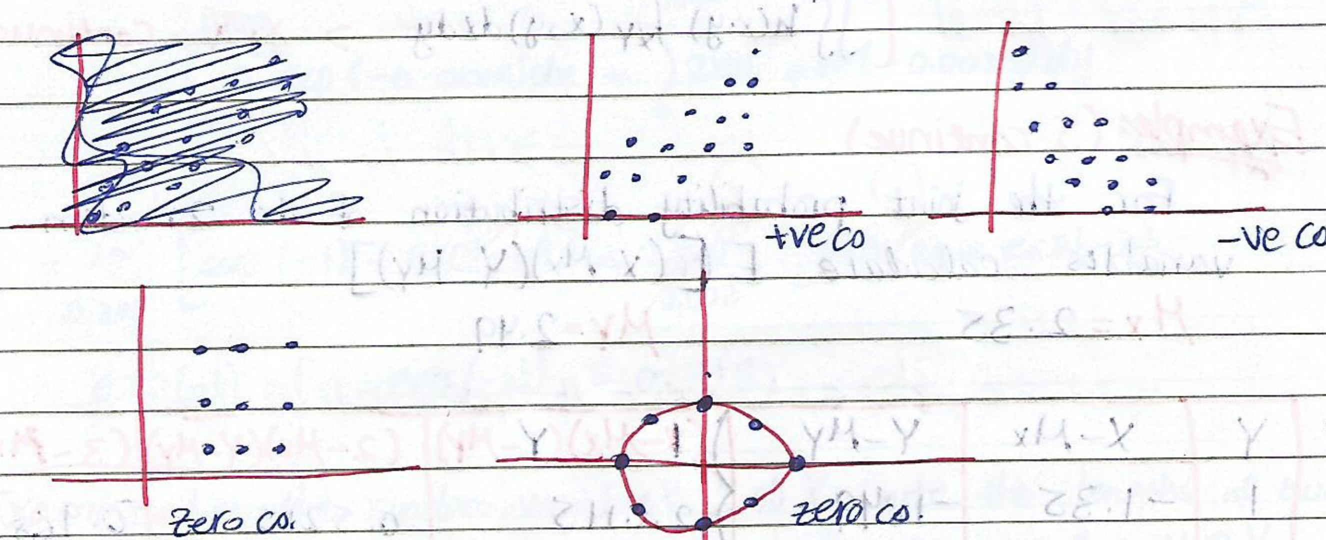
Covariance: $Cov(X,Y)$ OR σ_{XY}

$$\sigma_{XY} = E((X-\mu_X)(Y-\mu_Y)) = E(XY) - \mu_X\mu_Y$$

Covariance: is a measure of linear relationship between the random variables

* if the relationship between random variables is non linear?
the covariance might not be sensitive to the relationship

** Assumes all points are equally likely:



Correlation: $(\rho_{XY}) = \dfrac{cov(X,Y)}{\sqrt{V(X)V(Y)}} = \dfrac{\sigma_{XY}}{\sigma_X \sigma_Y}$

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

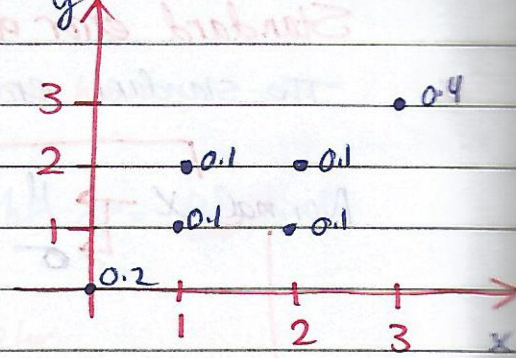
correlation: is a dimensionless quantity that can be used to compare the linear relationships between pairs of variables in different units.

* 2 random variables with non zero correlation are said to be correlated.

** the correlation is a measure of the linear relationship between random variables

Examples

Determine σ_{XY}, ρ_{XY}



$$\sigma_{XY} = \underbrace{E(XY)}_{\mu_{XY}} - \underbrace{E(X)}_{\mu_X}\,\underbrace{E(Y)}_{\mu_Y} \quad (1)$$

$E(XY) = \sum\sum XY\, f_{XY}(x,y)$ → discrete

$= 0*0*0.2 + 1*1*0.1 + 1*2*0.1$

$+ 2*1*0.1 + 2*2*0.1 + 0.4*3*3 = 0.1 + 0.2 + 0.2 + 0.4 + 3.6$

$\boxed{E(XY) = 4.5}$

مجموع $f_{XY}(x,y)$ ← عنا $x=1$, $x=2$

$E(X) = \sum x f_{XY}(x,y) = 0*0.2 + 0.2*1 + 0.2*2 + 3*0.4$

$= 0 + 0.2 + 0.4 + 1.2 = 1.8$

$E(Y) = \sum Y f_{XY}(x,y) = 0*0.2 + 1*0.2 + 2*0.2 + 3*0.4 = 1.8$

$\boxed{\sigma_{XY} = 4.5 - 1.8*1.8 = 4.5 - 3.24 = 1.26}$

(2) $\rho_{XY} = \dfrac{\sigma_{XY}}{\sigma_X \sigma_Y}$

$\sigma_X^2 = \sum (X - E(X))^2 f_{XY}(x,y) = (0-1.8)^2*0.2 + (1-1.8)^2*0.2 + (2-1.8)^2*0.2$

$+ (3-1.8)^2*0.4 = 0.648 + 0.128 + 0.008 + 0.576$

$\boxed{\sigma_X = 1.36}$

$\sigma_Y^2 = \sum (Y - \mu_y)^2 f_{XY}(x,y) = (0-1.8)^2 0.2 + (1-1.8)^2*0.2 + (2-1.8)^2*0.2 +$

$+ (3-1.8)^2*0.4 = 1.36$

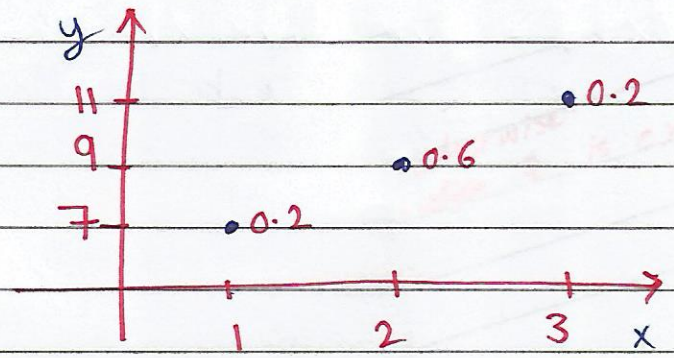
$\rho_{XY} = \dfrac{1.26}{\sqrt{1.36}\sqrt{1.36}} = 0.926$

Examples

Suppose that the random variable X has the following distribution: $P(X=1)=0.2$, $P(X=2)=0.6$, $P(X=3)=0.2$. Let $Y=2X+5$

That is $P(Y=7)=0.2$ $P(Y=11)=0.2$.

Determine the correlation between X & Y.



حسب المعادلة $Y=2X+5$

$X=1 \Rightarrow Y=7$

$X=2 \Rightarrow Y=9$

$X=3 \Rightarrow Y=11$

Ch7: Point Estimation of Parameters & Sampling distributions

§7.1: Point Estemation

Sample distribution: the probability distribution of a statistic

* Statistic ~~static~~ is a random variable, such as: mean, variance, ...

* the objective of point estimation: to select a single number based on sample data that is the most plausible value for θ.

* point estimate: numerical value $\hat{\theta}$ of a sample ~~static will be used~~ statistic

$\hat{\Theta}$

$\Rightarrow$ $\hat{\theta}$ = point estimate (numerical value) قيمة رقمية بتعطي تنبؤ عن

$\hat{\Theta}$ = point estimator) (sample statistic $\rightarrow h(X_1, X_2, \dots)$ القيمة الأصلية

المأخوذة من القيم الكبيرة

§7.2: Sampling Distributions & the Central Limit Theorem.

* The random variables $X_1, X_2, \dots, X_n$ are a random sample of size n if:

(a) the X_i's are independent random variables

(b) every X_i's has the samp probability distribution

* the primary purpose in taking arandom sample is to obtain information about the unknown population parameters.

* Statistic: any function of the observations in a random sample.

* Sampling of a statistic depends on apopulation in:

1- distribution

2- the size of the sample

3- the method of sample selection.

Sampling distribution of the mean.

$$\bar{X}\text{(point estimator)} = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

$$\mu_{\bar{x}}\text{ (normal distribution)} = \frac{\mu + \mu + \mu + \cdots + \mu}{n} = \mu$$

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2 + \sigma^2 + \cdots + \sigma^2}{n} = \frac{\sigma^2}{n} \Rightarrow \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$

Example: An electronics company manufactures resistors that have a mean resistance of $\overset{\mu}{100}$ ohms & a standard deviation of $\overset{\sigma}{10}$ ohms. The distribution of resistance is normal. Find the probability that a random sample of $\boxed{n=25}$ resistors will have an average resistance of fewer than 95 ohms. $P(\bar{X} < 95)$

$$P(\bar{x} < 95) = P\left(Z < \frac{95-100}{\frac{10}{\sqrt{25}}}\right) = P\left(Z < \frac{-5}{2}\right) = P(Z < -2.5)$$

$$= 0.0062 \text{ (rare event)}$$

Central limit theorem

If $X_1, X_2, \ldots, X_n$ is a random sample of size n taken from a population (finite/infinite) with mean μ and finite variance σ^2 & if $\bar{X}$ is a sample mean, the limiting form of the distribution of

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$$ is the standard normal distribution

* $n > 30$ → the central limit theorem will always apply.

$$\mu_{\bar{X}_1 - \bar{X}_2} = \mu_{\bar{X}_1} - \mu_{\bar{X}_2} = \mu_1 - \mu_2$$

$$\sigma^2_{\bar{X}_1 - \bar{X}_2} = \sigma^2_{\bar{X}_1} + \sigma^2_{\bar{X}_2} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}$$

If we have 2 independent populations with means μ_1 & μ_2 and variances σ_1^2 & σ_2^2 and if $\bar{X}_1$ & $\bar{X}_2$ are the sample means of 2 independent random samples of size n_1 & n_2 from these populations, then the sampling distribution of

$$Z = \frac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

is approximately standard normal if the conditions of the central limit theorem apply.

* If 2 populations are normal, the sampling distribution of Z is exactly standard normal.

Example: random variables.

$\bar{X}_1$ is normal with $\mu_1 = 5000$ $\sigma_1 = 40$ $n_1 = 16$

$\bar{X}_2$ is normal with $\mu_2 = 5050$ $\sigma_2 = 30$ $n_2 = 25$

What is the probability that the difference in the 2 samples means $\bar{X}_2 - \bar{X}_1$ is at least 25?

$$P(\bar{X}_2 - \bar{X}_1 \geq 25) = \cancel{P\left(\frac{\bar{X}_2 - \bar{X}_1 - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n} + \sigma_2^2}}\right)}$$

$$= P\left(\frac{\bar{X}_2 - \bar{X}_1 - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \geq \frac{25 - (5000 - 5050)}{\sqrt{\frac{(40)^2}{16} + \frac{(30)^2}{25}}}\right)$$

$$P\left(Z \geq \frac{-25}{\sqrt{136}}\right) = P(Z \geq -2.14) = 0.9838$$

§ 7-3.3: Standard error & Reporting a Point Estimate

The Standard error of an estimator $\hat{\Theta}$ is its standard deviation given by $\sigma_{\hat{\Theta}} = \sqrt{V(\hat{\Theta})}$. If the standard error involves unknown parameters that can be estimated, substitution of those values into $\sigma_{\hat{\Theta}}$ produces an estimated standard error, denoted by $\hat{\sigma}_{\hat{\Theta}}$ $se(\hat{\Theta})$

بيستخدموا لما بعرف σ للPopulation

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} \rightarrow \text{Standard error of } \bar{X}$$

$$SE(\bar{X}) = \hat{\sigma}_{\bar{X}} = \frac{S}{\sqrt{n}} \rightarrow \text{estimated standard error of } \bar{X}$$

بستخدمها لما ما بعرف σ لـ Population

بيكون عندي S لـ sample من Population

Example: Data:- (sample)

41.60 41.48 42.34 41.95 41.86
42.18 41.72 42.26 41.81 42.04

(1) $\bar{X} = \frac{\sum x}{n} = \frac{419.24}{10} = 41.924 \rightarrow$ sample mean

(2) $\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$ σ unknown for population

But $S = 0.284$

$SE(\bar{X}) = \hat{\sigma}_{\bar{x}} = \frac{S}{\sqrt{n}} = \frac{0.284}{\sqrt{10}} = 0.0898$

Ch 8: Statistical Intervals for a Single Sample

§8.1.1: Development of the confidence interval & Its Basic properties.

confidence interval: interval of the form: ~~the~~

$P\{L \le \mu \le U\} = 1-\alpha$

$L \rightarrow$ lower-confidence limit $U \rightarrow$ Upper-confidence limit

$1-\alpha \rightarrow$ confidence coefficient where $0 \le \alpha \le 1$

standardize

$$P\left(-z_{\frac{\alpha}{2}} \le \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \le z_{\frac{\alpha}{2}}\right) = 1-\alpha$$

ترتيب

$$P\underbrace{\left\{\bar{X} - z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{X} + z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{2}}\right\}}_{\text{random interval}} = 1-\alpha$$

If $\bar{x}$ is the ~~rand~~ sample mean of a random sample of size n from a normal population with known variance σ^2, a $100(1-\alpha)\%$ CI (confidence interval) on μ is given by

$$\bar{x} - z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{x} + z_{\frac{\alpha}{2}}\frac{\sigma}{\sqrt{n}}$$

where $z_{\frac{\alpha}{2}}$ is the upper $\frac{100\alpha}{2}\%$ of the standard normal distribution

Examples:-

64.1 64.7 64.5 64.6 64.5 64.3 64.6 64.8
64.2 64.3

$\sigma = 1$

Find 95% CI

$$\bar{x} = \frac{\sum x}{n} = \frac{644.6}{10} = 64.46$$

$n = 10 \qquad \sigma = 1$

$95\% = 100(1-\alpha)\%$

$1 - \alpha = 0.95$

$\Rightarrow \alpha = 0.05$

$Z_{\frac{\alpha}{2}} = Z_{\frac{0.05}{2}} = Z_{0.025} = 1.96 \Rightarrow$ بروح على جدول Z

وبدور على قيمة Z اللي الاحتمالية عندها (0.025) وباخذها موجبة دائماً

$$\bar{x} - Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}} \leq \mu \leq \bar{x} + Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}$$

$$64.46 - 1.96 * \frac{1}{\sqrt{10}} \leq \mu \leq 64.46 + 1.96 * \frac{1}{\sqrt{10}}$$

$$63.84 \leq \mu \leq 65.08$$