



تقدم لجنة

ملخص لمادة:

إحصاء و احتمالات

جزيل الشكر للطالب:

حمزة اسماعيل



* Probability and statistics :-

* Ch1:- Introduction to statistics :-

* Relationship between Probability and inferential statistics :-

1. Population :- المجتمع الإحصائي
 ⇒ contains all elements.

2. Sample :- العينة الإحصائية
 ⇒ Part of population.

* Data Analysis :-

Measures	Population	Sample
1- <u>Mean</u> المتوسط الحسابي	$m = \frac{\sum X}{n}$	$\bar{x} = \frac{\sum Y}{n}$
2- <u>Median</u> المتوسط	med متوسط القيم	med متوسط القيم
3- <u>Mode</u> المتوال	الرقم الأكثر تكراراً	الرقم الأكثر تكراراً
4- <u>Rang</u> رتب	$R = \max - \min$	$R = \max - \min$
5- <u>Variance</u> التباين	$\sigma^2 = \frac{\sum (X - \bar{X})^2}{n}$	$S^2 = \frac{\sum (X - \bar{X})^2}{n-1}$
6- <u>standard deviation</u>	$\sigma = \sqrt{\frac{\sum (X - \bar{X})^2}{n}}$	$S = \sqrt{\frac{\sum (X - \bar{X})^2}{n-1}}$

* Graphical Diagnostic :-

1. scatter plot :- الرسم البياني

2. stem and leaf Plot.

- ↳ single stem and leaf Plot.
- ↳ double stem and leaf Plot.

3- Histogram Plot.

* mean of Histogram = $\frac{\sum (\text{midPoint} * \text{freq})}{\sum F}$
 = $\sum (\text{midPoint} * \text{Rel Fre})$

* Relative freq = $\frac{\text{freq}}{\sum F}$.

* Graphical Diagnostic :-

↳ Box plot :-

First Quartile (Q_1) and third Quartile (Q_3)

① location of $Q_1 = 0.25(n+1)$.

② location of $Q_2 = 0.5(n+1)$.

③ location of $Q_3 = 0.75(n+1)$.

$Q_1 = 2.75$
 $Q_2 = 8.75$

① Value of $Q_1 = \text{Value of location (2)} + 0.75 (\text{U. loc (3)} - \text{U. loc (2)})$.

② Value of $Q_3 = \text{Value of location (8)} + 0.25 (\text{U. loc (9)} - \text{U. loc (8)})$.

* Inter-Quartile Rang (IQR) :-

$IQR = Q_3 - Q_1$

lower out lier = $Q_1 - 1.5 IQR$.

upper out lier = $Q_3 + 1.5 IQR$.

* Ch2:- Probability :-

* Sample space (S) :- الفضاء العيني

⇒ The set of all possible outcomes.

* The Union (U) and the Intersection (∩)

- Union (U) (الاتحاد) ⇒ The event that contain all elements of the group (A) and group (B)

- Intersection (∩) (التقاطع) ⇒ العناصر المشتركة

- Mutually exclusive (disjoint) :- الحوادث المتباعدة

$A \cap B = \emptyset$. A, B haven't element in common

* Venn diagram :- الشكل فين

$(A \cap \bar{A}) = \bar{A}$ تقاطع

$(A \cup \bar{A}) = A$ اتحاد

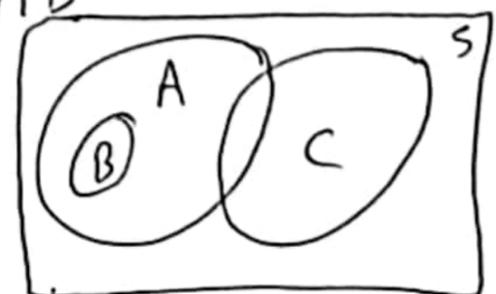
$(A \cup A') = S$ اتحاد $A' \rightarrow A$ complement

$S' = \bar{S}$. $\bar{\bar{S}} = S$.

$(A')' = A$.

$(A \cap B)' = A' \cup B'$

$(A \cup B)' = A' \cap B'$.



* Probability and statistics :-

* Ch2 :- Probability :-

* A Permutation :- التباديل

an arrangement of all or part of a set of objects.

1- Linear Permutation :-

[1] The number of permutation of (n) object is $n!$.

[2] The number of permutation of (n) distinct object taken (r) at a time is $P_r = \frac{n!}{(n-r)!}$.

[3] The number of permutation of (n) object where (n_1) repeated items, (n_2) repeated items:

$$P = \frac{n!}{n_1! n_2! \dots n_k!}$$

2- Circular Permutation :-

[1] The number of permutation of (n) objects arranged in a circle is $(n-1)!$:- without reference point.

[2] The number of permutation of (n) objects with reference point $n!$.

* Combinations :-

The number of combinations of (n) distinct object taken (r) at a time is

$$\binom{n}{r} = \frac{n!}{r! (n-r)!}$$

* arranged \Rightarrow Permutation

distinguish \Rightarrow Permutation.

and \Rightarrow Intersection تقاطع

or \Rightarrow Union \cup

both \Rightarrow Intersection تقاطع

* The probability :-

A probability of an event (A) is the sum of the weight of all sample points in (A) $\Rightarrow 0 \leq P(A) \leq 1$.

A_1, A_2, A_3, \dots mutually exclusive events

$$P(A_1 \cup A_2 \cup A_3 \dots) = P(A_1) + P(A_2) + P(A_3)$$

* Theorems :-

1- the probability of event (A) is $P(A) = \frac{n}{N}$ \rightarrow number of events in (A)
 $N \rightarrow$ number of sample space

2- If (A) and (B) are two events then $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

3- If (A) and (B) mutually exclusive $P(A \cup B) = P(A) + P(B)$.

4- If A_1, A_2, A_3 - mutually exclusive $P(A_1 \cup A_2 \cup A_3 \dots) = P(A_1) + P(A_2) + P(A_3)$

5- for three events A, B and C :- $P(A \cup B \cup C) = P(A) + P(B) + P(C) + P(A \cap B \cap C) - P(A \cap B) - P(A \cap C) - P(B \cap C)$.

6- If (A) and (A') are complementary event $P(A) + P(A') = 1$
 $(A \cup A') = S \Rightarrow P(S) = 1$.

7- The conditional probability of (B) given (A) denoted by $P(B|A)$.
 $P(B|A) = \frac{P(A \cap B)}{P(A)}$.

8- Two event (A), (B) are independent. $P(B|A) = P(B)$. $P(A|B) = P(A)$.

9- the event (A) and (B) both occur $P(A \cap B) = P(A) P(B|A)$.
 $P(A \cap B) = P(A) * P(B)$.

* Probability and statistics :-

* Ch 3 :- Random Variable and Probability distributions :-

* Random Variable = المتغير العشوائي
is a function that associates a real number with each element in the sample space. (S).

* Discrete sample space and the continuous sample space :-

Discrete sample space → Table

Continuous sample space → function

* Discrete Random Variable :-

Probability mass function or Probability of the discrete random variable if

1. $f(x) \geq 0$. 2. $\sum_x f(x) = 1$.

3. $P(X=x) = f(x)$.

* The Cumulative distribution function of a discrete random variable :-

$$F(x) = P(X \leq x) = \sum_{t \leq x} f(t) \quad -\infty < x < \infty$$

* Continuous Probability Distributions :-

Typical density function

$$P(a < X < b) = \int_a^b f(x) dx$$

* Continuous Probability Distributions

Probability density function for the continuous random variable (X) if.

1. $f(x) \geq 0$ 2. $\int_{-\infty}^{\infty} f(x) dx = 1$

3. $P(a < X < b) = \int_a^b f(x) dx$.

* The Cumulative distribution function of a continuous random variable with density function $f(x)$ is :-

$$F(x) = P(X \leq x) = \int_{-\infty}^x f(x) dx \quad -\infty < x < \infty$$

$$P(a < X < b) = F(b) - F(a)$$

* Joint Probability distribution or Probability mass function of the

discrete random variable (X) and (Y) if

1. $f(x,y) \geq 0$ 2. $\sum_x \sum_y f(x,y) = 1$

3. $P(X=x, Y=y) = f(x,y)$

* Joint density function of the continuous random variable (X) and (Y) if.

1. $f(x,y) \geq 0$ 2. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x,y) dx dy = 1$

3. $P((X,Y) \in A) = \iint_A f(x,y) dx dy$.

* The marginal distributions of (X, Y) :-

- marginal distributions for the discrete

$$g(x) = \sum_y f(x,y) \quad h(y) = \sum_x f(x,y)$$

- marginal distributions for continuous

$$g(x) = \int_{-\infty}^{\infty} f(x,y) dy \quad h(y) = \int_{-\infty}^{\infty} f(x,y) dx$$

* The Conditional distribution of (X, Y) :-
conditional distribution of (Y) given that (X)

$$f(y|x) = \frac{f(x,y)}{g(x)} \quad g(x) > 0$$

conditional distribution of (X) given that (Y)

$$f(x|y) = \frac{f(x,y)}{h(y)}$$

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

* Probability and statistics :-

* Ch 4 :- Mathematical Expectation :-

* Let X, Y be a random variable with Probability distribution $f(x)$

Then the Mean or expected Value (of X)

if (X) is discrete $M = E(X) = \sum X f(x)$

if (X) is continuous $M = E(X) = \int_{-\infty}^{\infty} X f(x) dx$

the mean of $g(x)$

if $g(x)$ is discrete $M = E(g(x)) = \sum g(x) f(x)$

if $g(x)$ is continuous $M = E(g(x)) = \int_{-\infty}^{\infty} g(x) f(x) dx$

the mean of Joint function of $g(x, y)$

if $g(x, y)$ is discrete $E(g(x, y)) = \sum_x \sum_y g(x, y) f(x, y)$

if $g(x, y)$ is continuous $E(g(x, y)) = \int_x \int_y g(x, y) f(x, y) dx dy$

* Let X, Y be a random variable with Probability distribution $f(x)$ and mean (M) the Variance of (X) is

if (X) is discrete $\sigma^2 = E[(X-M)^2] = \sum (X-M)^2 f(x)$

if X is continuous $\sigma^2 = E[(X-M)^2] = \int_{-\infty}^{\infty} (X-M)^2 f(x) dx$

$\sigma_x^2 = E(X^2) - M_x^2 \Rightarrow E(X^2) = \sum X^2 f(x) \quad M = \sum X f(x)$

$\sigma_x^2 = E(X^2) - M_x^2 \Rightarrow E(X^2) = \int X^2 f(x) \quad M = \int X f(x)$

* Means of Linear combination

$$M(ax+by+c) = aM_x + bM_y + c.$$

$$M(2x-1) = 2M_x - 1.$$

Variance of linear combination

$$\sigma_{(ax+by+c)}^2 = a^2 \sigma_x^2 + b^2 \sigma_y^2 + 2ab \sigma_{xy}.$$

$$\sigma_{(ax+by)}^2 = a^2 \sigma_x^2 + b^2 \sigma_y^2$$

$$\sigma_{(ax-by)}^2 = a^2 \sigma_x^2 + b^2 \sigma_y^2$$

* Variance of $g(x)$

$$\sigma_{g(x)}^2 = E(g(x)^2) - M_{g(x)}^2.$$

discrete $\Rightarrow E(g(x)^2) = \sum g(x)^2 f(x)$

continuous $\Rightarrow E(g(x)^2) = \int g(x)^2 f(x) dx.$

* Covariance of (X, Y) :

$$\sigma_{xy} = E(XY) - M_x M_y.$$

discrete $\Rightarrow E(XY) = \sum_x \sum_y XY f(x, y)$

continuous $\Rightarrow E(XY) = \int_x \int_y XY f(x, y) dx dy.$

if X, Y are independent then

$$E(XY) = E_x E_y.$$

$$\Rightarrow \sigma_{xy} = 0.$$

* Correlation coefficients :-

$$r_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$$

$$\sigma_{xy} = r_{xy} \sigma_x \sigma_y.$$

$$E(XY) - M_x M_y = r_{xy} \sigma_x \sigma_y.$$

$$M_x = \int g(x) f(x, y).$$

$$M_y = \int h(y) f(x, y).$$

* Probability and statistics :-

* Ch5: Discrete Uniform Distribution :-

* Binomial Distribution :-

خوارزميات :- مجموع احتمالات = واحد .

$$B(x, n, p) = \binom{n}{x} p^x q^{n-x}$$

$$p + q = 1 \Rightarrow \boxed{q = 1 - p}$$

x ← العود المطلوب إيجاد p ← عدد مرات تكرار التجربة .
 p ← احتمالية وقوع الحدث مرة واحدة .

$$B(x=a) = B(a) - B(a-1)$$

$$B(x \leq a) = B(a)$$

$$B(x < a) = B(a-1)$$

$$B(x \geq a) = 1 - B(a-1)$$

$$B(x > a) = 1 - B(a)$$

$$P(1 \leq x < 5) = B(4) - B(0)$$

1, 2, 3, 4

$$P(2 < x < 9) = B(8) - B(2)$$

3, 4, 5, 6, 7, 8

$$\text{Mean} = np \quad \sigma^2 = npq$$

ملاحظات :-

n ← أكبر رقم بالسؤال p ← $(0 \leq p \leq 1)$

أقل $\Rightarrow \leq$, at most $\Rightarrow \geq$, أكبر $\Rightarrow \geq$, at least

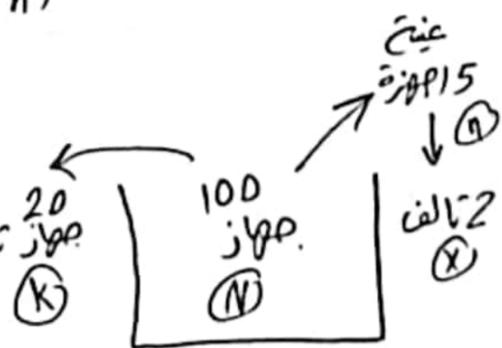
* Hyper geometric :-

فيها أربع أقسام :-

$$H(x, N, n, k) = \frac{\binom{k}{x} \binom{N-k}{n-x}}{\binom{N}{n}} \quad \begin{matrix} n < N \\ x < k \end{matrix}$$

$$\text{mean} = \frac{nk}{N}$$

$$\sigma^2 = \frac{N-n}{N-1} \times n \times \frac{k}{N} \times (1 - \frac{k}{N})$$



* ملاحظات :-

N ← العينة الكبيرة (أكبر رقم)

n ← العينة المأخوذة من العينة الصغيرة الكبيرة

k ← عدد ال defective بالعينة الكبيرة

x ← احتمال وجود ال defect في العينة الصغيرة

* حالة خاصة وهي :-

$$\boxed{\frac{n}{N} \leq 0.05} \rightarrow \text{Binomial Distribution}$$

Hyper \rightarrow Binomial

$$N = 1000$$

$$n = 5$$

$$x = 2$$

$$k = 100$$

$$n = 5$$

$$x = 2$$

$$p = \frac{100}{1000} = 0.1$$

* Negative and geometric distribution :-

$$b(x, k, p) = \binom{x-1}{k-1} p^k q^{x-k}$$

k ← متى اولى مرة تحققت فيه التجربة نجحها

x ← احتمال النجاح من (n) تجارب

$$\text{mean} = \frac{k}{p} \quad \sigma^2 = \frac{kq}{p}$$

* geometric distribution :-

هو حالة خاصة من ال negative يكون فيها $k=1$

$$b(x, 1, p) = \binom{x-1}{1-1} p^1 q^{x-1}$$

$$g(x, p) = p q^{x-1}$$

$$p + q = 1 \Rightarrow \boxed{q = 1 - p}$$

* Poisson Distribution :-

$$P(x, \lambda) = \frac{e^{-\lambda t} (\lambda t)^x}{x!}$$

$x!$

نستخدم جداول ال Poisson بنفس طريقة ال Binom.

$$P(x=a) = P(a) - P(a-1) \quad * P(2 < x < 6)$$

$$P(x \geq a) = 1 - P(a-1) \quad = P(5) - P(2)$$

$$P(x \leq a) = P(a) \quad * P(3 \leq x \leq 7)$$

$$P(x > a) = 1 - P(a) \quad = P(7) - P(2)$$

$$P(x < a) = P(a-1)$$