SPI **STATISTICS** IEW FROM CA ER
C

Brigida A. Roscom

 \sim $^{-1}$

rearn Calculus First applies to all mathematical sci
While no previous knowledge of statistics is assumed
in calculus is a requisite for students to work their plies to all mathematic eie
sui
as It is a sumed in calculus is a requisite for students to work their own
ackground in calculus is a requisite for students to work their owr
ough the courses in mathematical statistics or the classica
I theories. A sound ba Exercise is a requisite for sources in mathematical solutions in calculated in calculate a good working through the courses in mathematical statistics or the cla he courses in mathematical statistics or
ies. A sound background in calculus that is
ics, includes a good working knowledge n
ba tat
cu
es heories. A sound background in calculus that is essential for atistics, includes a good working knowledge of multiple partial differentiation and power series.
Ser is a tour of statistics via calculus with the aim of provi bund background in calculus that is esse statistics, includes a good working knowledge of mu

i, partial differentiation and power series.

aper is a tour of statistics via calculus with the aim of prov
 ook at statistics In the preparation of this tour, the au integration, partial differentiation and power series.

here is a tour of statistics via calculus with the aim of pro-
be at statistics. In the preparation of this tour, the author This paper is a tour of statistics via calculus with the aim of providir *eper look at statistics*. In the preparation of this tour, the author religity on existing literature on the subject of mathematical statistics. a calculus with
aration of this
bject of mather look at statistics. In the preparation of th
n existing literature on the subject of mat
stended view on mean, median, and m arti
r is
isti
isti ifferentiation and power series.

our of statistics via calculus with the d
vi
epsu

n existing literature on the subject of mathematical statistics.
 An Existing Statistics are not the subject of mode. The three averages

and mode, are popularly seen as single numerical ati:
ee
e ev
a
a
c **n mean, median, and mode.** The three ave
mode, are popularly seen as single num
the features of a data set usually a sample he he for a seen as sin

and the fortune of a data get upuelly median, and mode, are popularly seen as single nun
h summarize the features of a data set, usually a sample
the computational concepts of these average n, median, and mode, are popularly seen as single numerical
hich summarize the features of a data set, usually a sample or a
pulation data. The computational concepts of these averages are
prward. But it is obvious that th e
ob
et by pulation data. The computational concepts of these averages are

forward. But it is obvious that the computations have limitations

e population data set is infinite. It is this limitation that compels

tend our view on But
on
vie
lity is obvious that the computations have limit ard. But it is obvious that the computations have limital pulation data set is infinite. It is this limitation that con
that compute that contains that contained by the set of location. The background forward. But it is obvious that the computations have limitation the population data set is infinite. It is this limitation that computed our view on these measures of location. The backgro hich this view is set up, is bui e population data set is infinite. It is
tend our view on these measures of
sich this view is set up, is built up extend our view on these measures of location. The bac i]
oo
ct his view is set up, is built upon the concepts of random
bability, distributions and expectations. To give you a feel
pts, consider the following definitions. y, distributions and expectations. To give you a
onsider the following definitions.
If X is a discrete random variable with dist
identify that the read variability, distr

epts, consider

extracted if your _id
ng
re ese concepts, consider the following define

the following defined

the Section of the

discrete random variable with disti
 $[X = x]$ as the probability that the rand

in the function $f(.)$ defined by defined by If X
 \therefore will \therefore
 \therefore if x

if x $x_2, ..., x_n, ...$
akes the valu

the *X* takes the value *x*, then the function *f*(.) define
\n
$$
f(x) = \begin{cases} P[X = x], \text{ if } x = x_j, j = 1, 2, ..., n, ... \\ 0, \text{ if } x \neq x_j \end{cases}
$$

 $\mathbf d$ is also called the *probability mass* tion or *pmf* of *X*.

fion or *pmf* of *X*. function or pmf of X .

 \emph{pmf} of X .
Efinition. a continuous random variable X w

on $f(.)$ such that 7.2 F
fu domain There exists a function $f(.)$ such that

$$
P(X \leq x) = \int_{-\infty}^{x} f(u) du,
$$

der
s th is called the *probability density function* or *pdf* of X.
a density function $f(x)$ has the following properties : called the *probability density fu*

ensity function $f(.)$ has the follo

%) has the following properties :
 $2, \dots, r$ $f(x) = 0$ for $x \neq x_j$ and \sum $f(x, y) > 0$ for $j = 1, 2, ..., r$ $f(x) = 0$ for $x \neq x_j$ and \sum

screte.
 $f(x) \ge 0$ and $\int_{-\infty}^{\infty} f(x)dx = 1$ if X is cont

continuous.

Expeditional

<u>Expedition</u> DEFINITION. The *mean*, μ_X , or *expected value*, $\xi[X]$ of a ariable X is defined by [ˈh
ef

from variable *X* is defined by

\n
$$
\xi[X] = \begin{cases}\n\sum_{j} x_{j} f(x_{j}), & \text{if } X \text{ is discrete with mass points at } x_{j}, \text{ and } \\
\int_{-\infty}^{+\infty} x f(x) dx, & \text{if } X \text{ is continuous with pdf } f(x).\n\end{cases}
$$
\nNote that the definition assumes that the series is absolutely convergent.

definition assumes that the series is at hat the definition assumes that the series is absoluted
or the integral exists, respectively; otherwise, the mean doe
 $\frac{1}{2}$
 \frac h
e r the integral exists, respectively; other

virtion. The *median* x_{med} of a random virtillowing inequalities: not exist.

EFINITION. The **median** x_{med} of a random variable X is a num the following inequalities: he
g i

satisfying the following inequalities: P[Xs xma]2 and P[X2 Xmed]

BRIGIDA A. RosCOM continuous random variable, then the median of X satisfying the state of X satisfying the state of \int_{0}^{∞}

is a continuous random variable, then
\n
$$
\int_{-\infty}^{x_{\text{med}}} f(x) dx = \frac{1}{2} = \int_{x_{\text{med}}}^{\infty} f(x) dx.
$$

enoted by x_{mo} , of a random variable X
erical value which satisfies $\frac{d}{dt}$ (x) is a numerical value which satis $\}$. m
S

$$
f(x_{\text{mo}}) = \max_{x} \{f(x)\}.
$$

 $f(x)$.
 x_{mo}) = max{ $f(x)$ }.

ode, if it exists, is that value of X which max

c).
Ug EXAMPLE. A random variable X, which is **normally dist** μ and variance σ^2 , where $\sigma > 0$, has pdf given by r
iar
E ean μ and variance σ^2 , where $\sigma > 0$, has pdf given by $\frac{2}{\epsilon}$

$$
f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right].
$$

 $\mathfrak{z},$

us show that the mean or expected value of X is
$$
\mu
$$
. By Def
\n
$$
\xi[X] = \int_{-\infty}^{+\infty} x \left(\frac{1}{\sigma \sqrt{2\pi}} \right) \exp \left[-\frac{1}{2} \left(\frac{x - \mu}{\sigma} \right)^2 \right] dx.
$$

the
bt the the state of the variables, i.e., by letting $z =$ \mathfrak{O}^{\dagger}

$$
\xi[X] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} (\sigma z + \mu) e^{\frac{-z^2}{2}} dz,
$$

$$
\xi[X] = \frac{\sigma}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} ze^{\frac{-z^2}{2}} dz + \frac{\mu}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{\frac{-z^2}{2}} dz
$$

$$
= \frac{\sigma}{\sqrt{2\pi}} A + \frac{\mu}{\sqrt{2\pi}} B,
$$

THE MINDANAO FORUM)
R

where

$$
A = \int_{-\infty}^{+\infty} \frac{-z^2}{2} dz = -e^{-\frac{z^2}{2}} \Bigg|_{-\infty}^{+\infty}
$$

= $\lim_{z \to +\infty} (-e^{-z^2/2}) - \lim_{z \to -\infty} (-e^{-z^2/2}) = 0,$

and

$$
\frac{1}{\sqrt{2\pi}} B = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{z^2/2} dx = 1,
$$

we have second integrand is the
dom variable with mean 0 and va
Combining the results for *A* and

–
ecc
iab ntegrand is the density function of that
h mean 0 and variance 1.
results for A and B, the mean of a normall iun
an normal ean 0 and var

random variable X is sults for A and B , the mean of a normally dist
 ar
bi
ar vit
|e
| i .
m

om variable X is
\n
$$
\xi[X] = \frac{1}{\sqrt{2\pi}} \cdot 0 + \mu \cdot 1 = \mu.
$$

XAMPLE. Again, let us consider finding the mode of a not
I random variable X. Invoking Definition 7.5, we shall app
of differentiation to obtain the maximum value of the pdf
he first derivative of $f(x)$ of Example 7.6 to ın
dif
fi mation *X*. Invoking Definition 7.5, we shat
ation to obtain the maximum value of the vative of $f(x)$ of Example 7.6 to zero, we .5
al
z pply the
df of X. f differentiation to obtain the maximum value of the pdf of X.
he first derivative of $f(x)$ of Example 7.6 to zero, we obtain $\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}^2$

ating the first derivative of
$$
f(x)
$$
 of Example 7.6 to zero, we obtain
\n
$$
f'(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2} \left(\frac{x-\mu}{\sigma}\right)^2} \cdot \cdot \cdot - \left(\frac{x-\mu}{\sigma}\right) = 0, \text{ or}
$$
\n
$$
e^{-\frac{1}{2} \left(\frac{x-\mu}{\sigma}\right)^2} \cdot \cdot \cdot - \left(\frac{x-\mu}{\sigma}\right) = 0.
$$

rst factor cannot be zero, so $\left(\frac{m-r}{\sigma}\right) = 0$, hence $x = \mu$. an
de e
V

cond derivative test would confirm if indeed the solution $x = \frac{1}{2}$
haximum value of $f(x)$. Now since the maximum value of $f(x)$. Now since alı
İ

$$
f''(x) = \frac{-1}{\sqrt{2\pi}\sigma} \left\{ \left(e^{-\frac{1}{2} \left(\frac{x-\mu}{\sigma} \right)^2} \right) \cdot \left(\frac{x-\mu}{\sigma} \right)^2 + \left(\frac{1}{\sigma} \right) \cdot \left(e^{-\frac{1}{2} \left(\frac{x-\mu}{\sigma} \right)^2} \right) \right\}
$$

$$
= \frac{-1}{\sqrt{2\pi}\sigma} \left(e^{-\frac{1}{2} \left(\frac{x-\mu}{\sigma} \right)^2} \right) \left[\left(\frac{x-\mu}{\sigma} \right)^2 + \left(\frac{1}{\sigma} \right) \right],
$$

mmediating have

$$
f''(\mu) = \left[-\frac{1}{\sqrt{2\pi}\sigma} \cdot 1 \right] \left[\frac{1}{\sigma} \right] < 0,
$$

we immediately have

$$
f''(\mu) = \left[-\frac{1}{\sqrt{2\pi}\sigma} \cdot 1 \right] \left[\frac{1}{\sigma} \right] < 0,
$$

> 0. Thus the maximum of *f* is attained at $x = \mu$, which shows that the $x_{\text{mo}} = \mu$. $\int f$ ode $x_{\text{mo}} = \mu$

and simple regression line. Consider a set of bivariate data.

1, 2, ..., $n\}$. A simple linear regression model postulates at elationship between the variables X and Y. The mode mple regression line. Consider a set of bivary ..., $n\}$. A simple linear regression model posonship between the variables X and Y. Th inear regression mode

e variables X and elationship between the variables *X* and *Y*. The most $Bx_i + \varepsilon_i$ expressed as Bx

$$
y_i = A + Bx_i + \varepsilon_i
$$

tended to express an approximately linear relation between X and Y , re any deviation from a perfect straight line relationship is attributed. led to express an approximately linear relation between λ
iy deviation from a perfect straight line relationship is attri
in random variation. In the model, x_i is some fixed value ne
 $\frac{1}{x_i}$

and n a
bu
f om a perfect straight line relationship is at
ation. In the model, x_i is some fixed value
primally distributed random variables with
in standard deviation σ . And for each x $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$ or random variation. In the model, x_i is some fixed value of λ
are independent normally distributed random variables with mear ndom variation. In the model, x_i is some fixed value of X ; the ıt
di
ar ori
wn
av e independent normally distributed random variables with mean zero
e same unknown standard deviation σ . And for each x_i , a set of
ng y_i values have mean $A + Bx_i$. The ε_i 's, where nd
1 (
Th ar
m $\frac{1}{2}$ deviation
deviation
deviation v_i values have mean $A + Bx_i$. The ε_i 's, whe

$$
\varepsilon_i = y_i - (A + Bx_i),
$$

^e deviations of the Y values from the regression line $Y =$
delow is a table of naire of height u and weight u from a

a table of pairs of height x and
i. The pairs (x,y) are plotted in values from the regression line $Y = A + BX$.
airs of height x and weight y from a sample of 15
(x,y) are plotted in Figure 1. it
devi
de female of pairs of height x and we

infors. The pairs (x,y) are plotted in Figure 133 Figure 1.

THE MINDANAO FORUM

	THE MINDANAO FORUM														
χ	50	158	155	55	56	157	57,	158	160	162	162	163	165	166	166
	104	100	112	98	105	95	14	108	99	14	95	110	100	18	106

Heights and weights of 15 female studes

Scatter Plot of Weight vs. Hei

Assumptions in a Regression Mod
134

BRIGIDA A.

parameters A and B are estimated using sample data. Bet
of these parameters is the *principle of least square* ar
para
set sir
cij
10 ata. Behind the
t *squares*. The If these parameters is the *principle of least squares*. The clares that when choosing among the possible lines to ivariate set of data, the line that best fits is that line which must as possible the sum S of the squared nese parameters is the *principle of leas* declares that when choosing among the possible lines to bivariate set of data, the line that best fits is that line which small as possible the sum S of the squared vertical among the possible line
hat best fits is that line \sqrt{S} of the squared ve represent to that, the line that best fits is that line wh small as possible the sum S of the squared vertifrom the points to the line. The mathematical procedure ivariate set of data, the line that best fits is that
mall as possible the sum S of the same et
in small as possible the sum S of the sq
om the points to the line. The mathematical is possible the sum S of the squared ve points to the line. The mathematical procedure uses partial differentiation. What follows is om the points to the line. The mathematical procedure of best line uses partial differentiation. What follows is the wards the least squares estimates for the regression line A and B . Let S be the sum of the squared I the points to the line. The mathematical procedure of

est line uses partial differentiation. What follows is the

rds the least causes estimates for the ro
lo[.]
res
1s, this best line uses partial differentiation. What follows is the towards the least squares estimates for the regression line ints *A* and *B*. Let *S* be the sum of the squared deviations, then Vh
th
le ROsCOM beta by observation in the set of squares es s A and B . Let S be the sum of or
ec In the sum of the squared deviations, then
 $\mathbf{a}_i - (A + Bx_i)$.

efficients *A* and *B*. Let *S* be t

$$
S = \sum_{i=1}^{n} \varepsilon_i = y_i - (A + Bx_i).
$$

 $\sum_{i=1}$

artial differentiation technique to minimize S, we provide the state of the state of $\sum_{i=1}^{n}$ (*V* and *PV*) and *n* with

loying partial differentiation technique,
\n
$$
\frac{\partial S}{\partial A} = -2 \sum_{i=1}^{n} (Y_i - A - BX_i) = 0, \text{ and}
$$
\n
$$
\frac{\partial S}{\partial B} = -2 \sum_{i=1}^{n} X_i (Y_i - A - BX_i) = 0.
$$

the estimates for *A* and *B* that minimizes *S*, denoted by *a* and ively, are the solutions to the equations below: or
ol s,
w re the solutions to the equations belo
 $-bX_i$) = 0, and

$$
\sum_{i=1}^{n} (Y_i - a - bX_i) = 0, \text{ and}
$$

$$
\sum_{i=1}^{n} X_i (Y_i - a - bX) = 0.
$$

p each of the equations above, we have $\sum_{n=0}^{n} x^{n} = 0$

$$
\sum_{i=1}^{n} Y_i - na - b \sum_{i=1}^{n} X_i = 0
$$

$$
\sum_{i=1}^{n} X_i Y_i - a \sum_{i=1}^{n} X_i - b \sum_{i=1}^{n} X_i^2 = 0, \text{ or}
$$

\n
$$
an + b \sum_{i=1}^{n} X_i = \sum_{i=1}^{n} Y_i, \text{ and}
$$

\n
$$
a \sum_{i=1}^{n} X_i + b \sum_{i=1}^{n} X_i^2 = \sum_{i=1}^{n} X_i Y_i.
$$

The last pair of equations above, are known as the normal equations. They have solutions for b and a as follows:

$$
b = \frac{\sum_{i=1}^{n} X_i Y_i - \frac{1}{n} \left[(\sum_{i=1}^{n} X_i)(\sum_{i=1}^{n} Y_i) \right]}{\sum_{i=1}^{n} X_i^2 - \frac{1}{n} (\sum_{i=1}^{n} X_i)^2} = \frac{\sum_{i=1}^{n} (X_i - \overline{X})(Y_i - \overline{Y})}{\sum_{i=1}^{n} (X_i - \overline{X})^2} \text{ and }
$$

$$
a = \overline{Y} - b\overline{X}.
$$

The Poisson distribution as a limiting distribution for the binomial distribution. Two of the commonly used discrete distribution models in the life and social sciences are the Poisson and Binomial distributions. The Poisson distribution provides a probabilistic model for a wide class of phenomena. Examples are the number of telephone calls during a given period of time, the number of particles emitted from a radioactive source, and the number of cars passing by an intersection point. The binomial distribution is by far the most important discrete distribution. An experiment or an activity follows a **binomial model** if it has n independent trials with two possible outcomes per trial: either a specific event occurs or does not occur. The probability p of the occurrence of the said event remains the same from trial to trial. Typical examples are flipping of a coin, getting a defective or nondefective product, and having a boy or a girl for a child.

BRIGIDA A. ROSCOM

.
be x;
s,
m p) and $P(x; \lambda)$ be the bind
graatively. For agah, fixed nass functions, resp
hat $\lambda = np$ remains
x; n, p) = P(x; np). nctions, respectively. For each fixed x, as $n \to \infty$ and
np remains constant,
 $= P(x; np)$. $\int \alpha \, dx = np \, ren$
 $\lim_{n \to \infty} \frac{p}{n}$ n

 $B(x; n, p) = P(x; np).$ ic)
1

Proof. $B(x; n, p)$ denotes the probability that an event *E* has a convention *B*(*x*; *n*, *p*) is defined by *n*) denotes the probability that an event *E* has *x* nurn trials. The event *E* has a probability *p* of occurrenc ction $B(x; n, p)$ is defined by o
C
C the event E h
 \therefore n, p) is def \mathbf{a} ial. The function $B(x; n, p)$ is defined by

0 such that
$$
\lambda = np
$$
 remains constant,
\n
$$
\lim_{n \to \infty} B(x; n, p) = P(x; np).
$$
\nProof. $B(x; n, p)$ denotes the probability that an event *E* has *x* number
\ncurrences in *n* trials. The event *E* has a probability *p* of occurrence in
\ntrial. The function $B(x; n, p)$ is defined by
\n
$$
B(x; n, p) = \binom{n}{p} p^x q^{n-x}
$$
\n
$$
= \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x}
$$
\n
$$
= \frac{n(n-1) \cdots (n-x+1)}{x!} p^x (1-p)^{n-x} \qquad (7.8.1)
$$

llow p to vary with n, so take $p = \frac{1}{n}$, $n \ge 1$, $\lambda > 0$. Thut

stituting
$$
p = \frac{n}{n}
$$
 in (7.8.1), we have,
\n
$$
B(x; n, p) = \frac{n (n-1) \cdots (n-x+1)}{x!} \left[\frac{\lambda}{n} \right]^x \left[1 - \frac{\lambda}{n} \right]^{n-x}
$$
\n
$$
= 1 \left[1 - \frac{1}{n} \right] \cdots \left[1 - \frac{x-1}{n} \right] \frac{\lambda^x}{x!} \left[1 - \frac{\lambda}{n} \right]^n \left[1 - \frac{\lambda}{n} \right]^{-x}
$$

 ∞ , while x and λ remain constant, we obtain the follow-
 $\begin{bmatrix} 1 \end{bmatrix} \begin{bmatrix} x-1 \end{bmatrix}$

$$
\lim_{n \to \infty} 1 \left[1 - \frac{1}{n} \right] \cdots \left[1 - \frac{x-1}{n} \right] = 1 \text{ and}
$$

$$
\lim_{n \to \infty} \left[1 - \frac{\lambda}{n} \right]^{-x} = 1.
$$

From the definition of the number e , we also have

$$
\lim_{n \to \infty} \left[1 - \frac{\lambda}{n} \right]^n = e^{-\lambda} \,. \tag{7.8.2}
$$

Therefore, under the given limiting conditions.

$$
\lim_{n\to\infty} B(x;n,p) = P(x;\lambda) = \frac{e^{-\lambda}\lambda^x}{x!}.
$$

The amazing central limit theorem. When a finite set of independent random variables $X_1, X_2, ..., X_n$ has a common distribution, the probability distribution for their mean \overline{X} is approximately normal for large *n*. The precise statement of this observation is one of the most celebrated theorems of mathematics, the so-called Central Limit Theorem or CLT.

The Central Limit Theorem implies that if the sample n is large and yet a small fraction of the population size N so that independence of $X_1, X_2, ..., X_n$ is reasonable, we can approximate the probabilities of the sample mean \overline{X} using the table of areas under the normal curve. A special case of this theorem is stated below.

7.9 THEOREM (CENTRAL LIMIT THEOREM). Let $f(.)$ be a density function with mean μ and variance σ^2 . Let $\overline{X_n}$ be the sample mean of a random sample of size n from $f(.)$. Let the random variable Z_n be defined by

$$
Z_n = \frac{\overline{X_n} - \xi(\overline{X_n})}{\sqrt{var(\overline{X_n})}} = \frac{\overline{X_n} - \mu}{\sigma/\sqrt{n}}.
$$

Then, the distribution of Z_n approaches the standard normal distribution as n tends to infinity.

BRIGIDA A. RosCOM

Discussion. The amazing part of the Central Limit Theorem is the fact that nothing is assumed about the form of the original density funtion. The importance of this theorem as far as applications are concerned, is the fact that the mean of a random sample of size n from any distribution with finite variance σ^2 and mean μ , is approximately distributed as a normal random variable with mean μ and variance σ^2/n . The proof of this theorem will make use of the concepts of moments, and moment gene rating function of a standard normal random variable.

The *moment generating function* or *mgf* of a random variable X is defined as

$$
m(t) = \xi \Big[e^{tX} \Big] = \int_{-\infty}^{+\infty} e^{tX} f(x) dx.
$$

If the mgf of X exists, then $m(t)$ is continuously differentiable at some neighborhood of the origin. Differentiating $m(t)$, r times with respect to t, we obtain

$$
\frac{d^r}{dt^r}m(t)=\int_{-\infty}^{+\infty}x^r e^{tx}f(x)dx.
$$

Letting $t \to 0$, we have

$$
\frac{d^r}{dt^r} m(0) = \int_{-\infty}^{+\infty} x^r f(x) dx = \mu_r.
$$

The number μ_r is called the *rth moment of f(x)*. Replacing e^{tX} by its series expansion in ξ $\left[e^{tX}\right]$, we obtain

$$
m(t) = \xi \left[1 + Xt + \frac{1}{2!} (Xt)^2 + \frac{1}{3!} (Xt)^3 + \dots \right]
$$

= 1 + $\mu_1 + \frac{1}{2!} t^2 \mu_2 + \frac{1}{3!} t^3 \mu_3 + \dots$

THE MINDANAO FORUM

$$
= \sum_{i=0}^{\infty} \frac{\mu_r}{i!} t^r.
$$

gf of a standard normal random variable is $m(t) = e^{\frac{1}{2}t}$
oment generating function of $Z = m$ (t) can be all

Exercise memoric generating function of Z_n , $m_{Z_n}(t)$ can be shown
the standard normal random variable, $m(t)$ as *n* becomes la
Theorem 7.9 is proved. f Z_n , $m_{Z_n}(t)$ can be shown to
riable, $m(t)$ as *n* becomes large the standard normal random variable, $m(t)$ as *n* becomes largeorem 7.9 is proved.
of *Theorem* 7.9. Using the independence of *X*, *X*, *X*, *N* External 1.9 is proved that the state of σ^2

 P is proved.
 Pm 7.9. Using t Theorem 7.9. Using the independence of $X_1, X_2, ..., X_n$, we obtain

Proof of Theorem 7.9. Using the independence of
$$
X_1, X_2, ..., X_n
$$
,
\nn
\n
$$
m_{Z_n}(t) = \xi \left[e^{tZ_n} \right] = \xi \left[\exp \left(t \frac{\overline{X} - \mu}{\sigma / \sqrt{n}} \right) \right] = \xi \left[\exp \left(\frac{t}{n} \sum_{i=1}^n \frac{X_i - \mu}{\sigma / \sqrt{n}} \right) \right]
$$
\n
$$
= \xi \left[\prod_{i=1}^n \exp \left(\frac{t}{n} \cdot \frac{X_i - \mu}{\sigma / \sqrt{n}} \right) \right] = \prod_{i=1}^n \xi \left[\exp \left(\frac{t}{\sqrt{n}} \cdot \frac{X_i - \mu}{\sigma} \right) \right].
$$

we let $Y_i = (X_i - \mu) / \sigma$, then $m_{Y_i}(t)$, the more of Y_i , is independent of i since all Y_i 's iven $m_{Y_i}(t)$, the moment gen
of *i* since all Y_i 's have the
 $m_{Y_i}(t)$. Then of Y_i , is independent of

in. Let $m_Y(t)$ denote each m_Y Y_i , is independent of *i* since all Y_i 's have the sar
let $m_Y(t)$ denote each $m_Y(t)$. Then $\gamma(t)$ denote each $m_{Y_i}(t)$. Then

olution. Let
$$
m_Y(t)
$$
 denote each $m_{Y_i}(t)$. Then

\n
$$
\prod_{i=1}^n \xi \left[\exp\left(\frac{t}{\sqrt{n}} \cdot \frac{X_i - \mu}{\sigma}\right) \right] = \prod_{i=1}^n \xi \left[\exp\left(\frac{t}{\sqrt{n}} \cdot Y_i\right) \right]
$$
\n
$$
= \prod_{i=1}^n m_{Y_i}\left(\frac{t}{\sqrt{n}}\right) = \left[m_Y\left(\frac{t}{\sqrt{n}}\right) \right]^n.
$$

Hence,

$$
m_{Z_n}(t) = \left[m_Y\left(\frac{t}{\sqrt{n}}\right)\right]^n.
$$

BRIGIDA A. ROSCOM

In derivative of $m_Y(t/\sqrt{n})$ evaluated at $t = 0$ gives us the rth nt about the mean of the density f() divided by $(\sigma_s/n)^r$ so we may f
ear
nsi $\frac{1}{\sqrt{2}}$ t
d l bout the mean of the density $f(.)$ divided by $(\sigma \sqrt{n})^r$, so vaylor expansion write its Taylor expansion

its Taylor expansion
\n
$$
m_Y \left(\frac{t}{\sqrt{n}} \right) = 1 + \frac{\mu_1}{\sigma} \frac{t}{\sqrt{n}} + \frac{1}{2!} \frac{\mu_2}{\sigma^2} \left(\frac{t}{\sqrt{n}} \right)^2 + \frac{1}{3!} \frac{\mu_3}{\sigma^3} \left(\frac{t}{\sqrt{n}} \right)^3 + \cdots
$$

 $\mu_1 = 0$ and $\mu_2 = \sigma^2$, this may be writ

$$
m_Y\left(\frac{t}{\sqrt{n}}\right) = 1 + \frac{1}{n}\left(\frac{1}{2}t^2 + \frac{1}{3!\sqrt{n}}\frac{\mu_3}{\sigma^3}t^3 + \frac{1}{4!n}\frac{\mu_4}{\sigma^4}t^4 + \cdots\right).
$$
 (7.9.1)

all that $\lim_{n \to \infty} (1 + \frac{u}{n})^n = e^u$, if *u* is constant. Thus, if *u* rep
on within the parenthesis in (7.9.1), then it fo h
th he
iat the
re, parenthesis in (7.9.1), then it follows the
use $\lim u = \frac{1}{2}t^2$. Therefore, we have $\lim_{n \to \infty} u = \frac{1}{2}t^2$. Therefore, we have er
n within the parenthesis in
 $= e^{\frac{1}{2}t^2}$, because $\lim_{n \to \infty} u = \frac{1}{2}t^2$ \int the \int , be

$$
\lim_{n\to\infty}m_{Z_n}(t)=\lim_{n\to\infty}[m_Y(\frac{t}{\sqrt{n}})]^n=\lim_{n\to\infty}(1+\frac{\mu}{n})^n=e^{\frac{1}{2}t^2},
$$

 $(1 + \frac{\mu}{n})^n = e^{\frac{1}{2}t^2}$,
ent generating function at, at the limit, Z_n has the same moment generating function as that of candard normal variable. the standard normal variable.

REFERENCES

- M
Iil [| Mood, A. M. and Graybill, to the Theory of $\frac{1}{10}$ ati
-
- Vilks, S.S. *Mathematical Statistics*, Wiley, 1962
 ati
.