කැලණිය විශ්වවිද්‍යාලය - ශ්‍රී ලංකාව

දුරස්ථ සහ අඛණ්ඩ අධ්‍යාපන කේන්ද්‍රය

වාණිජ හා කළමනාකරණ අධ්‍යයන පීඨය

ව්‍යාපාර කළමනාකරණවේදී (සාමාන්‍ය) උපාධි දෙවන පරීක්ෂණය (බාහිර) - 2012

2015 ජූනි/ ජූලි

BMGT E 2045 - කළමනාකරණය සඳහා සංඛ්‍යානය

ප්‍රශ්න සංඛ්‍යාව : 08 කාලය : පැය 03 යි

ඕනෑම ප්‍රශ්න පහකට (05) පිළිතුරු සපයන්න.

සූත්‍ර සහ සංඛ්‍යාන වගු සපයා ඇත.

(01) අ) "සංඛ්‍යානය යනු දත්ත එකතු කිරීම, විශ්ලේෂණය කිරීම, නිර්ණය කිරීම, ඉදිරිපත් කිරීම සහ සංවිධානය කිරීම පිළිබඳ විද්‍යාවයි." ව්‍යාපාර සංඛ්‍යානයේ නිර්වචන පිළිබඳව අවධානය යොමු කරමින් ඉහත ප්‍රකාශය පැහැදිලි කරන්න.

(ලකුණු 04)

ආ) ව්‍යාපාර ක්ෂේත්‍රයට සංඛ්‍යානයේ ඇති වැදගත්කම ඉස්මතු කරන්න.

(ලකුණු 06)

ඇ) පහත සඳහන් දෑ කෙටියෙන් දක්වන්න.

(i) ප්‍රාථමික දත්ත

(ii) සමීක්ෂණ

(iii) අසමූහිත සංඛ්‍යාත ව්‍යාප්තිය

(iv) ගණිතමය සීමා

(v) විවික්ත විචල්‍යය (ලකුණු 10)

(මුළු ලකුණු 20)

(02) ආ) විස්තරාත්මක දත්ත විශ්ලේෂණය කිරීම සඳහා යොදා ගනු ලබන ක්‍රම මොනවාද? පැහැදිලි කරන්න.

(ලකුණු 06)

ආ) පසුගිය කාල පරිච්ඡේදය තුළ සේවකයන් දෙදෙනෙකුගේ එක සමාන ක්‍රියා වන්හි ප්‍රතිඵල පහතින් දක්වා ඇත.

	සේවකයා A	සේවකයා B
ක්‍රියාව නිම කිරීමේ මධ්‍යනය කාලය (මිනිත්තු)	30	25
සම්මත අපගමනය (මිනිත්තු)	06	04

(i) මෙම ක්‍රියාව සම්පූර්ණ කිරීමේදී වැඩි ඒකාකාරී බවක් දක්වන්නේ කුමන සේවකයා ද? පැහැදිලි කරන්න.

(ii) ක්‍රියාව ඉතා ඉක්මණින් සම්පූර්ණ කරන්නේ කුමන සේවකයාද? පැහැදිලි කරන්න.

(ලකුණු 06)

ඇ) Y විචල්‍යයට ඇත්තේ අගයන් 5 ක් පමණක් බව සහ ඒවාට අදාල සාපේක්ෂ සංඛ්‍යාති අගයන් පහත පරිදි බව සිතන්න.

Y	සාපේක්ෂ සංඛ්‍යාතිය	සමුච්චිත සාපේක්ෂ සංඛ්‍යානය
0	5/100	
1	25/100	
2	30/100	
3	25/100	
4	15/100	

(i) සම්මුච්චිත සාපේක්ෂ සංඛ්‍යාතය සම්පූර්ණ කරන්න.

(ii) පංති පළල ඒකිය වන පරිදි පංති 05 ගන්න. Y හි පංති අගයන්හි මධ්‍ය අගය හඳුනාගන්න. මෙම දත්ත වලට සාපේක්ෂ සංඛ්‍යාත ඡාලරේඛය ගොඩනගන්න.

(iii) ඔබගේ සාපේක්ෂ සංඛ්‍යාත ඡාල රේඛය භාවිතයෙන් මධ්‍යනයය, මධ්‍යස්ථය හා මාතය දක්වන්න.

(iv) සාපේක්ෂ සංඛ්‍යාතය සම්භාවිතාවක් ලෙස නිර්නය කර ඇත්නම් Y හි අගය 2 ට සමාන හෝ වැඩිවීමේ සම්භාවිතාවය කොපමණද?

(ලකුණු 08)

(මුළු ලකුණු 20)

(03) අ) සම්භාවිතා ප්‍රවේශයන් කෙටියෙන් පැහැදිලි කරන්න.

(ලකුණු 04)

ආ) ගිණුමේ ආයු කාලය සහ ගිණුමේ ශේෂය මත ජාතික බැංකුව සිය ගිණුම් කාණ්ඩ ගත කර ඇත.

ශේෂය / ගිණුමේ ආයු කාලය	Rs. 0 - 999	Rs. 1000 - 4999	Rs. 5000 හෝ ඊට වැඩි
වසර 3 ට අඩු	700	100	400
වසර 3 හෝ ඊට වැඩි	200	400	200

ගිණුම් 2000 ක කාණ්ඩයකින් සසම්භාවී ලෙස එක් ගිණුමක් තෝරා ගත හොත් එය,

(i) ගිණුමේ ශේෂය රු. 5000 හෝ ඊට වැඩි සහ ආයුකාලය වසර 3 ට අඩු වීමේ සම්භාවිතාව, සොයන්න.

(ii) ගිණුමේ ශේෂය රු. 0 - 999 හෝ ආයු කාලය වසර 3 හෝ ඊට වැඩි වීමේ සම්භාවිතාව සොයන්න.

(iii) ගිණුමේ ශේෂය රු. 1000 හෝ ඊට වැඩි වීමේ සම්භාවිතාව සොයන්න.

(iv) ගිණුමේ ආයුකාලය වසර 3 ට වඩා අඩුබව දී ඇත්නම්, ගිණුමේ ශේෂය රු. 1000 ට අඩු වීමේ අසම්භාව්‍ය සම්භාවිතාව කොපමණද?

(v) මෙම බැංකුවේ ගිණුම් ශේෂයන් සහ ගිණුමේ ආයු කාලයන් එකිනෙකට ස්වායත්ත වේද? එසේ වන්නේ හෝ නොවන්නේ මන්ද?

(vi) මෙම බැංකුවෙන් ගිණුම් 14 ක් සසම්භාවීව තෝරාගත්තේ යැයි සිතන්න. මෙහි G යනු, අවම වශයෙන් ගිණුම් 5 ක් වත් වසර තුනකට අඩු ආයුකාලයකින් යුතුවේ."
එනම් G සහ එහි අනුපූරකය කුමක්ද?

(ලකුණු 12)

ඇ) සිසුවකු සංඛ්‍යානය පරීක්ෂණයෙන් සමත්වීමේ සම්භාවිතාව 2/3 ක් සහ ඔහු සංඛ්‍යානය සහ ගණිතය යන පරීක්ෂණ දෙකින්ම සමත් වීමේ සම්භාවිතාවය 4/15 කි. ඔහු අවම

වශයෙන් එක් පරීක්ෂණයකින් සමත්වීමේ සම්භාවිතාව 4/5 කි. ඔහු ගණිතය පරීක්ෂණයෙන් සමත්වීමේ සම්භාවිතාවය කුමක්ද?

(ලකුණු 04)

(මුළු ලකුණු 20)

(04) ඇ) ප්‍රමත ව්‍යාප්තියක ආවේණික ලක්ෂණ දක්වන්න.

(ලකුණු 04)

ආ) නිශ්චිත ආර්ථික කාණ්ඩයක පවුල් හි මාසික මධ්‍යන්‍ය ආහාර වියදම රු. 1300 සහ සම්මත අපගමනය රු. 200 ක් වන ලෙස ප්‍රමතව ව්‍යාප්ත වී ඇත.

(i) වියදම රු. 900 ට වඩා අඩු වීමේ සමානුපාතය කුමක්ද?

(ii) වියදම රු. 1000 සහ රු. 1200 ත් අතර වීමේ ප්‍රතිශතය කුමක්ද?

(iii) වියදම රු. 1200 ට වඩා අඩු වීමේ හෝ රු. 1500 ට වඩා වැඩි වීමේ ප්‍රතිශතය කුමක්ද?

(iv) ඉහළම වියදම් දරණ 14%, දෙනාගේ අවම වියදම් අගය කොපමණද?

(ලකුණු 12)

ඇ) කිසියම් විභාගයක විකල්ප 04 ක් සහිතව බහුවරණ ප්‍රශ්න 25 ක් ඇත.

(i) අවම වශයෙන් නිවැරදි පිළිතුරු 8 ක්

(ii) උපරිම ලෙස නිවැරදි පිළිතුරු 12 ක්

(iii) හරියටම නිවැරදි පිළිතුරු 06 ක්, ලැබීමේ සම්භාවිතාවය සොයන්න.

(ලකුණු 04)

(මුළු ලකුණු 20)

(05) අ) පහත සඳහන් දෑ කෙටියෙන් දක්වන්න.

(i) සංගහනය.

(ii) පරාමිති

(iii) පංගු නියදීම (Quota sampling)

(iv) පහසු නියදීම

(v) මධ්‍ය සීමා ප්‍රමේය

(ලකුණු 05)

ආ) එක්තරා නගරයක වෙසෙන පුද්ගලයන්ගෙන් 20% ක් ප්‍රමාණයක් මාසික වැටුප ලෙස උපයන මුදල රු. 200,000 ට වඩා වැඩිය. මෙම නගරයෙන් පුද්ගලයන් 100 ක නියදියක් සසම්භාවීව තෝරා ගෙන පරීක්ෂා කළ විට එම නියදියේ පහත සඳහන් සමානුපාතයන් කෙසේ විය හැකිද?

(i) රු. 200,000 වඩා වැඩි මාසික වැටුප් හිමි පුද්ගල සංඛ්‍යාව 36% වඩා වැඩිවීම.

(ii) රු. 200,000 ට වැඩි මාසික වැටුප් හිමි පුද්ගල සංඛ්‍යාව 29% ට වඩා අඩු වීම.

(iii) රු. 200,000 ට වැඩි මාසික වැටුප් හිමි පුද්ගල සංඛ්‍යාව 25% සහ 35% අතර වීම.

(ලකුණු 15)

(06) අ) සාමාන්‍ය සමාජීය සමීක්ෂණයක අසන ලද, "ඔබ ගත යුතු නියමිත ආහාර ප්‍රමාණයට වඩා වැඩියෙන් ආහාර අනුභව කරන්නේද?" ගැටළුවට මිනිසුන් 680 නියැදියකින් ලද ප්‍රතිචාර පහත වගුවෙන් සහ ප්‍රතිඵලයෙන් ඉදිරිපත් කර ඇත.

	ඔව්	නැත	එකතුව
පිරිමි	151	177	328
ගැහැණු	92	260	352
එකතුව	243	437	680

ප්‍රතිඵලය

> වගු අගයන්

ඔව් නැත

පිරිමි 151 177

ගැහැණු 92 260

> පරීක්ෂණය < කයි වර්ග (.Table, correct = FALSE)

> . Test

පියර්සන් කයි වර්ග පරීක්ෂණය

දත්ත - වගුව

X - වර්ගය = 29.2788, df = 1, p - අගය = 6.268e - 08

(i) නියමිත ප්‍රමාණයට වඩා වැඩියෙන් ආහාර ගන්නා පිරිමින්ගේ ප්‍රතිශතය කුමක්ද?

(ii) නියමිත ප්‍රමාණයට වඩා වැඩියෙන් ආහාර ගන්නා ගැහැණුන්ගේ ප්‍රතිශතය කුමක්ද?

(iii) පිරිමින්ට සාපේක්ෂව තමා වැඩියෙන් ආහාර ගන්නවා යැයි සිතන ගැහැණුන්ගේ ප්‍රතිශතය කුමක්ද?

(iv) පිරිමින්ට සාපේක්ෂව තමා වැඩියෙන් ආහාර ගන්නවා යැයි සිතන ගැහැණුන්ගේ අනුපාතය කුමක්ද?

(v) ගතයුතු ප්‍රමාණයට වඩා වැඩියෙන් ආහාර ලබා ගැනීමේ ගැහැණුන් සහ පිරිමින්ගේ සමානුපාතයේ වෙනසක් ඇති බවට ප්‍රමාණවත් සාක්ෂි ඇත්ද? P අගය වාර්තා කර එය අර්ථකථනය කරන්න.

(ලකුණු 14)

ආ) කීර්තිය හා අධ්‍යාපන අතර ඇති සම්බන්ධතාවය පිළිබඳව රැකියා 102 ඇසුරින් ලබා ගන්නා ලද තොරතුරු පහත ප්‍රතිපායන විශ්ලේෂණ ප්‍රතිඵලය මගින් දක්වා ඇත.

ප්‍රතිඵලය

call :

Im (formula = prestige ~ education, data = prestige)

Residuals : (අවශේෂ)

Min (අවම)	1q	මධ්‍යනාස්‍ය	3Q	උපරිම (Max)
-26.0397	-6.5228	0.6611	6.7430	18.1636

සංගුණකයන් :

	Estimate Std.	Error t value Pr(> |t|)		
(Intercept)	-10.732	3.677	-2.919	0.00434 **
අධ්‍යාපනය	5.361	0.332	16.148	<2e-16 ***
signif. codes :	0'***'	0.001 '**'	0.01 '*'	0.05 '.' 0'1 " ' 1

Residual standard error : 9.103 on 100 degrees of freedom (සුවලන අංකය)
(අවශේෂ සම්මත දෝෂය)

Multiple R - squared : 0.7226, Adjected R - squared : 0.72

F - Statistic : 260.8 on 1 and 1000 DF,

P - Value : < 2.2e - 16

(i) අධ්‍යාපනය (x) කීර්තියේ (y) ශ්‍රිතයක් ලෙස ගත් විට ප්‍රතිපායන සමීකරණය කුමක්ද?

(ii) අධ්‍යාපනය සහ කීර්තිය අතර ඇති සම්බන්ධතාවය දැක්වීමට ඇති සාක්ෂි ඔබ විස්තර කරන්නේ කෙසේද?

(iii) ආදායම සහ කීර්තිනාමය අතර ඇති සහ සම්බන්ධතාවය කුමක්ද?

(ලකුණු 06)
(මුළු ලකුණු 20)

(07) ඇ) එක්තරා ඖෂධ වර්ගයක් බඳුන්වලට ඇසිරීමේදී එහි නියමිත බර විය යුත්තේ අවුන්ස 20 කි. එය නිශ්චිත තත්ත්වයෙන් යුතු නිෂ්පාදනයක් සේ සලකනු ලබන්නේ එසේ නම් පමණි. මෙම ඖෂධය, බඳුන්වලට අසුරණ නිෂ්පාදන මාර්ගයෙන් සසම්භාවී බඳුන් 60 ක් ගෙන පරීක්ෂා කළ විට එහි සාමාන්‍ය බර අවුන්ස 20.32 බවද, සම්මත් අපගමනය අවුන්ස 0.6 ක් බවද හෙළි විය. $\alpha = 0.01$ විට මෙම නිෂ්පාදන ක්‍රියාවලිය නියමිත තත්ත්වයෙන් බැහැර වී ඇත්ද යන්න නිගමනය කරන්න.

(ලකුණු 10)

ආ) එක් දාදු කැටයක් 1000 වාරයක් පෙරලූ විට "6" ලැබුණ වාර ගණන 204 කි. මෙම දාදු කැටය සාධාරණ නොවන බව පිළිගැනීමට සාක්ෂි ප්‍රමාණවත් වේද? ($\alpha = 0.05$)

(ලකුණු 10)

(මුළු ලකුණු 20)

(08) ඇ) පොත් ප්‍රකාශක සමාගමක සංස්කාරවරයෙකුගේ ප්‍රකාශය වූයේ, පාඨ ග්‍රන්ථයක් ලිවීමට ගතවන සාමාන්‍ය කාලය උපරිම වශයෙන් මාස 15 ක් වන බවය. පාඨ ග්‍රන්ථ කතෘන් 16 ක් සසම්භාවී නියැදියක් මගින් හෙළි වූයේ පාඨ ග්‍රන්ථයක් ලිවීමට ගතවන සාමාන්‍ය කාලය මාස 12.5 ක් බවය. එහි සම්මත අපගමනය මාස 3.6 වන පරිදි ප්‍රමතව ව්‍යාප්ත වී ඇත්නම් 0.025 ක වෙසෙසියා මට්ටමකට අනුව සංස්කාරකවරයාගේ ප්‍රකාශයේ සත්‍යතාවය පරීක්ෂා කරන්න.

(ලකුණු 10)

ආ) ශාරීරික ව්‍යායාම කිරීම මගින් තරබාරුකම අඩුකර ගැනීමට වඩා ආහාර පාලනයෙන් එය සිදුවේද යන්න තීරණය කිරීමට නියැදි සමීක්ෂණයක් කරන ලදී. එහි තොරතුරු පහත සඳහන් වේ. ආහාර පාලනය පමණක් කරන නියැදියේ නියැදි ප්‍රමාණය 42 ක් සහ බර අඩුවීමේ නියැදියේ මධ්‍යනය කිලෝ ග්‍රෑම් 5.9 ක්ද, නියැදි සම්මත අපගමනය කිලෝ ග්‍රෑම් 4.1 ක් ද විය. ශාරීරික ව්‍යායාම පමණක් කරන නියැදියේ ප්‍රමාණය 47 ක් වූ අතර බර අඩුවීමේ නියැදි මධ්‍යනය කිලෝ ග්‍රෑම් 4.1 ක් සහ නියැදි සම්මත අපගමනය කිලෝ ග්‍රෑම් 3.7 ක් විය . ඉහත තොරතුරු මත තරබාරුකම අඩුකර ගැනීම සඳහා ශාරීරික ව්‍යායාම වලට වඩා ආහාර පාලනය සුදුසුද යන්න ($\alpha = 0.05$) පරීක්ෂා කරන්න.

(ලකුණු 10)

(මුළු ලකුණු 20)

FORMULAE

Summary Measures

Sample Mean

$$\bar{x} = \frac{x_1 + x_2 + \cdots + x_n}{n} = \frac{\sum x_i}{n}$$

Sample Standard Deviation

$$s = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}} = \sqrt{\frac{\sum x_i^2 - n\bar{x}^2}{n-1}}$$

Probability Rules

- **Complement rule**

  $P(A^C) = 1 - P(A)$

- **Addition rule**

  General: $P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$

  For independent events:

  $P(A \text{ or } B) = P(A) + P(B) - P(A)P(B)$

  For mutually exclusive events: $P(A \text{ or } B) = P(A) + P(B)$

- **Multiplication rule**

  General: $P(A \text{ and } B) = P(A)P(B \mid A)$

  For independent events: $P(A \text{ and } B) = P(A)P(B)$

  For mutually exclusive events: $P(A \text{ and } B) = 0$

- **Conditional Probability**

  General: $P(A \mid B) = \dfrac{P(A \text{ and } B)}{P(B)}$

  For independent events: $P(A \mid B) = P(A)$

  For mutually exclusive events: $P(A \mid B) = 0$

Discrete Random Variables

Mean

$$E(X) = \mu = \sum x_i p_i = x_1 p_1 + x_2 p_2 + \cdots + x_k p_k$$

Standard Deviation

$$s.d.(X) = \sigma = \sqrt{\sum (x_i - \mu)^2 p_i} = \sqrt{\sum \left(x_i^2 p_i\right) - \mu^2}$$

Binomial Random Variables

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

where $\dbinom{n}{k} = \dfrac{n!}{k!(n-k)!}$

Mean

$$E(X) = \mu_X = np$$

Standard Deviation

$$s.d.(X) = \sigma_X = \sqrt{np(1-p)}$$

Normal Random Variables

- $z\text{-score} = \dfrac{\text{observation} - \text{mean}}{\text{standard deviation}} = \dfrac{x - \mu}{\sigma}$
- Percentile: $x = z\sigma + \mu$
- If X has the $N(\mu, \sigma)$ distribution, then the variable

  $Z = \dfrac{X - \mu}{\sigma}$ has the $N(0,1)$ distribution.

Normal Approximation to the Binomial Distribution

If X has the $B(n, p)$ distribution and the sample size n is large enough (namely $np \geq 10$ and $n(1-p) \geq 10$),

then X is approximately $N\left(np, \sqrt{np(1-p)}\right)$.

Sample Proportions

$$\hat{p} = \frac{x}{n}$$

Mean

$$E(\hat{p}) = \mu_{\hat{p}} = p$$

Standard Deviation

$$\text{s.d.}(\hat{p}) = \sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}$$

Sampling Distribution of $\hat{p}$

If the sample size n is large enough (namely, $np \geq 10$ and $n(1-p) \geq 10$)

then $\hat{p}$ is *approximately* $N\left(p, \sqrt{\dfrac{p(1-p)}{n}}\right)$.

Sample Means

Mean

$$E(\bar{X}) = \mu_{\bar{X}} = \mu$$

Standard Deviation

$$s.d.(\bar{X}) = \sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$$

Sampling Distribution of $\bar{X}$

If X has the $N(\mu, \sigma)$ distribution, then $\bar{X}$ is

$$N(\mu_{\bar{X}}, \sigma_{\bar{X}}) \Leftrightarrow N\left(\mu, \frac{\sigma}{\sqrt{n}}\right).$$

If X follows *any* distribution with mean μ and standard deviation σ and n is large,

then $\bar{X}$ is *approximately* $N\left(\mu, \dfrac{\sigma}{\sqrt{n}}\right)$.

This last result is **Central Limit Theorem**

Population Proportion

Parameter p

Statistic $\hat{p}$

Standard Error

$$\text{s.e.}(\hat{p}) = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

Confidence Interval

$$\hat{p} \pm z^{*}\text{s.e.}(\hat{p})$$

Conservative Confidence Interval

$$\hat{p} \pm \frac{z^{*}}{2\sqrt{n}}$$

Large-Sample z-Test

$$z = \frac{\hat{p} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}$$

Sample Size

$$n = \left(\frac{z^{*}}{2m}\right)^2$$

Two Population Proportions

Parameter $p_1 - p_2$

Statistic $\hat{p}_1 - \hat{p}_2$

Standard Error

$$\text{s.e.}(\hat{p}_1 - \hat{p}_2) = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

Confidence Interval

$$(\hat{p}_1 - \hat{p}_2) \pm z^{*}\text{s.e.}(\hat{p}_1 - \hat{p}_2)$$

Large-Sample z-Test

$$z = \frac{\hat{p}_1 - \hat{p}_2}{\sqrt{\hat{p}(1-\hat{p})\left(\frac{1}{n_1} + \frac{1}{n_2}\right)}}$$

where $\hat{p} = \dfrac{n_1\hat{p}_1 + n_2\hat{p}_2}{n_1 + n_2}$

Population Mean

Parameter μ

Statistic $\bar{x}$

Standard Error

$$\text{s.e.}(\bar{x}) = \frac{s}{\sqrt{n}}$$

Confidence Interval

$$\bar{x} \pm t^{*}\text{s.e.}(\bar{x}) \qquad \text{df} = n-1$$

Paired Confidence Interval

$$\bar{d} \pm t^{*}\text{s.e.}(\bar{d}) \qquad \text{df} = n-1$$

One-Sample t-Test

$$t = \frac{\bar{x} - \mu_0}{\text{s.e.}(\bar{x})} = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} \qquad \text{df} = n-1$$

Paired t-Test

$$t = \frac{\bar{d} - 0}{\text{s.e.}(\bar{d})} = \frac{\bar{d}}{s_d/\sqrt{n}} \qquad \text{df} = n-1$$

Two Population Means

General

Parameter $\mu_1 - \mu_2$

Statistic $\bar{x}_1 - \bar{x}_2$

Standard Error

$$\text{s.e.}(\bar{x}_1 - \bar{x}_2) = \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

Confidence Interval

$$(\bar{x}_1 - \bar{x}_2) \pm t^{*}(\text{s.e.}(\bar{x}_1 - \bar{x}_2)) \qquad \text{df} = \min(n_1 - 1, n_2 - 1)$$

Two-Sample t-Test

$$t = \frac{\bar{x}_1 - \bar{x}_2 - 0}{\text{s.e.}(\bar{x}_1 - \bar{x}_2)} = \frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \qquad \text{df} = \min(n_1 - 1, n_2 - 1)$$

Pooled

Parameter $\mu_1 - \mu_2$

Statistic $\bar{x}_1 - \bar{x}_2$

Standard Error

$$\text{pooled s.e.}(\bar{x}_1 - \bar{x}_2) = s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

where $s_p = \sqrt{\dfrac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}}$

Confidence Interval

$$(\bar{x}_1 - \bar{x}_2) \pm t^{*}(\text{pooled s.e.}(\bar{x}_1 - \bar{x}_2)) \qquad \text{df} = n_1 + n_2 - 2$$

Pooled Two-Sample t-Test

$$t = \frac{\bar{x}_1 - \bar{x}_2 - 0}{\text{pooled s.e.}(\bar{x}_1 - \bar{x}_2)} = \frac{\bar{x}_1 - \bar{x}_2}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \qquad \text{df} = n_1 + n_2 - 2$$

One-Way ANOVA

$$\text{SS Groups} = \text{SSG} = \sum_{\text{groups}} n_i(\bar{x}_i - \bar{x})^2$$

$$\text{MS Groups} = \text{MSG} = \frac{\text{SSG}}{k-1}$$

$$\text{SS Error} = \text{SSE} = \sum_{\text{groups}} (n_i - 1)\, s_i^2$$

$$\text{MS Error} = \text{MSE} = s_p^2 = \frac{\text{SSE}}{N-k}$$

$$\text{SS Total} = \text{SSTO} = \sum_{\text{values}} (x_{ij} - \bar{x})^2$$

$$F = \frac{\text{MS Groups}}{\text{MS Error}}$$

ANOVA Table

Source	SS	DF	MS	F
Groups	SS Groups	$k-1$	MS Groups	F
Error	SS Error	$N-k$	MS Error	
Total	SSTO	$N-1$		

Confidence Interval $\bar{x}_i \pm t^{*}\dfrac{s_p}{\sqrt{n_i}}$ $\quad$ $\text{df} = N-k$

Under H_0, the F statistic follows an $F(k-1, N-k)$ distribution.

Regression

Linear Regression Model

Population Version:

Mean: $\mu_Y(x) = E(Y) = \beta_0 + \beta_1 x$

Individual: $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$

where ε_i is $N(0, \sigma)$

Sample Version:

Mean: $\hat{y} = b_0 + b_1 x$

Individual: $y_i = b_0 + b_1 x_i + e_i$

Standard Error of the Sample Slope

$$\text{s.e.}(b_1) = \frac{s}{\sqrt{S_{XX}}} = \frac{s}{\sqrt{\sum (x-\bar{x})^2}}$$

Confidence Interval for β_1

$b_1 \pm t^* \text{s.e.}(b_1)$ df $= n-2$

***t*-Test for β_1**

To test $H_0 : \beta_1 = 0$

$t = \dfrac{b_1 - 0}{\text{s.e.}(b_1)}$ df $= n-2$

or $F = \dfrac{MSREG}{MSE}$ df $= 1, n-2$

Parameter Estimators

$$b_1 = \frac{S_{XY}}{S_{XX}} = \frac{\sum (x-\bar{x})(y-\bar{y})}{\sum (x-\bar{x})^2} = \frac{\sum (x-\bar{x})y}{\sum (x-\bar{x})^2}$$

$$b_0 = \bar{y} - b_1 \bar{x}$$

Confidence Interval for the Mean Response

$\hat{y} \pm t^* \text{s.e.(fit)}$ df $= n-2$

where $\text{s.e.(fit)} = s\sqrt{\dfrac{1}{n} + \dfrac{(x-\bar{x})^2}{S_{XX}}}$

Residuals

$e = y - \hat{y} = \text{observed } y - \text{predicted } y$

Prediction Interval for an Individual Response

$\hat{y} \pm t^* \text{s.e.(pred)}$ df $= n-2$

where $\text{s.e.(pred)} = \sqrt{s^2 + (\text{s.e.(fit)})^2}$

Correlation and its square

$$r = \frac{S_{XY}}{\sqrt{S_{XX} S_{YY}}}$$

$$r^2 = \frac{SSTO - SSE}{SSTO} = \frac{SSREG}{SSTO}$$

where $SSTO = S_{YY} = \sum (y-\bar{y})^2$

Standard Error of the Sample Intercept

$$\text{s.e.}(b_0) = s\sqrt{\frac{1}{n} + \frac{\bar{x}^2}{S_{XX}}}$$

Confidence Interval for β_0

$b_0 \pm t^* \text{s.e.}(b_0)$ df $= n-2$

Estimate of σ

$$s = \sqrt{MSE} = \sqrt{\frac{SSE}{n-2}} \quad \text{where } SSE = \sum (y-\hat{y})^2 = \sum e^2$$

***t*-Test for β_0**

To test $H_0 : \beta_0 = 0$

$t = \dfrac{b_0 - 0}{\text{s.e.}(b_0)}$ df $= n-2$

Chi-Square Tests

Test of Independence & Test of Homogeneity

Expected Count

$$E = \text{expected} = \frac{\text{row total} \times \text{column total}}{\text{total } n}$$

Test Statistic

$$X^2 = \sum \frac{(O-E)^2}{E} = \sum \frac{(\text{observed} - \text{expected})^2}{\text{expected}}$$

df $= (r-1)(c-1)$

Test for Goodness of Fit

Expected Count

$E_i = \text{expected} = np_{i0}$

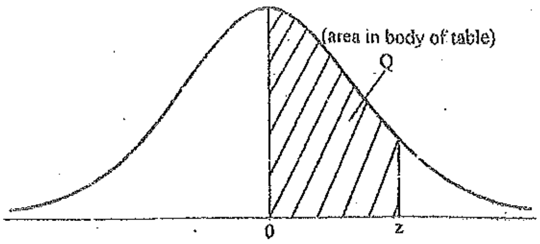
Test Statistic

$$X^2 = \sum \frac{(O-E)^2}{E} = \sum \frac{(\text{observed} - \text{expected})^2}{\text{expected}}$$

df $= k-1$

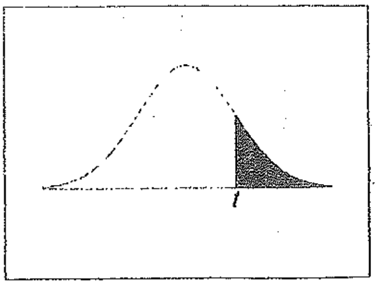
If Y follows a $\chi^2(df)$ distribution, then $E(Y) = df$ and $Var(Y) = 2(df)$.

Table 1 AREAS UNDER THE STANDARD NORMAL CURVE



z	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.0000	0.0040	0.0080	0.0120	0.0160	0.0199	0.0239	0.0279	0.0319	0.0359
0.1	0.0398	0.0438	0.0478	0.0517	0.0557	0.0596	0.0636	0.0675	0.0714	0.0753
0.2	0.0793	0.0832	0.0871	0.0910	0.0948	0.0987	0.1026	0.1064	0.1103	0.1141
0.3	0.1179	0.1217	0.1255	0.1293	0.1331	0.1368	0.1406	0.1443	0.1480	0.1517
0.4	0.1554	0.1591	0.1628	0.1664	0.1700	0.1736	0.1772	0.1808	0.1844	0.1879
0.5	0.1915	0.1950	0.1985	0.2019	0.2054	0.2088	0.2123	0.2157	0.2190	0.2224
0.6	0.2257	0.2291	0.2324	0.2357	0.2389	0.2422	0.2454	0.2486	0.2517	0.2549
0.7	0.2580	0.2611	0.2642	0.2673	0.2704	0.2734	0.2764	0.2794	0.2823	0.2852
0.8	0.2881	0.2910	0.2939	0.2967	0.2995	0.3023	0.3051	0.3078	0.3106	0.3133
0.9	0.3159	0.3186	0.3212	0.3238	0.3264	0.3289	0.3315	0.3340	0.3365	0.3389
1.0	0.3413	0.3438	0.3461	0.3485	0.3508	0.3531	0.3554	0.3577	0.3599	0.3621
1.1	0.3643	0.3665	0.3686	0.3708	0.3729	0.3749	0.3770	0.3790	0.3810	0.3830
1.2	0.3849	0.3869	0.3888	0.3907	0.3925	0.3944	0.3962	0.3980	0.3997	0.4015
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	0.4131	0.4147	0.4162	0.4177
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	0.4279	0.4292	0.4306	0.4319
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	0.4406	0.4418	0.4429	0.4441
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	0.4515	0.4525	0.4535	0.4545
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	0.4608	0.4616	0.4625	0.4633
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	0.4686	0.4693	0.4699	0.4706
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	0.4750	0.4756	0.4761	0.4767
2.0	0.4772	0.4778	0.4783	0.4788	0.4793	0.4798	0.4803	0.4808	0.4812	0.4817
2.1	0.4821	0.4826	0.4830	0.4834	0.4838	0.4842	0.4846	0.4850	0.4854	0.4857
2.2	0.4861	0.4864	0.4868	0.4871	0.4875	0.4878	0.4881	0.4884	0.4887	0.4890
2.3	0.4893	0.4896	0.4898	0.4901	0.4904	0.4906	0.4909	0.4911	0.4913	0.4916
2.4	0.4918	0.4920	0.4922	0.4925	0.4927	0.4929	0.4931	0.4932	0.4934	0.4936
2.5	0.4938	0.4940	0.4941	0.4943	0.4945	0.4946	0.4948	0.4949	0.4951	0.4952
2.6	0.4953	0.4955	0.4956	0.4957	0.4959	0.4960	0.4961	0.4962	0.4963	0.4964
2.7	0.4965	0.4966	0.4967	0.4968	0.4969	0.4970	0.4971	0.4972	0.4973	0.4974
2.8	0.4974	0.4975	0.4976	0.4977	0.4977	0.4978	0.4979	0.4979	0.4980	0.4981
2.9	0.4981	0.4982	0.4982	0.4983	0.4984	0.4984	0.4985	0.4985	0.4986	0.4986
3.0	0.4987	0.4987	0.4987	0.4988	0.4988	0.4989	0.4989	0.4989	0.4990	0.4990
3.1	0.4990	0.4991	0.4991	0.4991	0.4992	0.4992	0.4992	0.4992	0.4993	0.4993
3.2	0.4993	0.4993	0.4994	0.4994	0.4994	0.4994	0.4994	0.4995	0.4995	0.4995
3.3	0.4995	0.4995	0.4995	0.4996	0.4996	0.4996	0.4996	0.4996	0.4996	0.4997
3.4	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4997	0.4998

t-Distribution Table

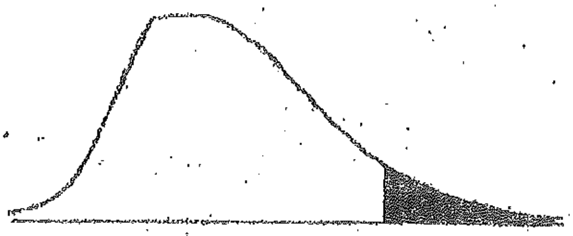


The shaded area is equal to α for $t = t_{\alpha}$.

df	$t_{.100}$	$t_{.050}$	$t_{.025}$	$t_{.010}$	$t_{.005}$
1	3.078	6.314	12.706	31.821	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.638	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
30	1.310	1.697	2.042	2.457	2.750
32	1.309	1.694	2.037	2.449	2.738
34	1.307	1.691	2.032	2.441	2.728
36	1.306	1.688	2.028	2.434	2.719
38	1.304	1.686	2.024	2.429	2.712
∞	1.282	1.645	1.960	2.326	2.576

CRITICAL VALUES OF CHI-SQUARE

This table contains the values of χ^2 that correspond to a specific right-tail area and specific numbers of degrees of freedom df.



χ^2

Possible values of χ^2

Degrees of Freedom *df*	Righ-tail Area 0.10	0.05	0.02	0.01
1	2.706	3.841	5.412	6.635
2	4.605	5.991	7.824	9.210
3	6.251	7.815	9.837	11.345
4	7.779	9.488	11.668	13.277
5	9.236	11.070	13.388	15.086
6	10.645	12.592	15.033	16.812
7	12.017	14.067	16.622	18.475
8	13.362	15.507	18.168	20.090
9	14.684	16.919	19.679	21.666
10	15.987	18.307	21.161	23.209
11	17.275	19.675	22.618	24.725
12	18.549	21.026	24.054	26.217
13	19.812	22.362	25.472	27.688
14	21.064	23.685	26.873	29.141
15	22.307	24.996	28.259	30.578
16	23.542	26.296	29.633	32.000
17	24.769	27.587	30.995	33.409
18	25.989	28.869	32.346	34.805
19	27.204	30.144	33.687	36.191
20	28.412	31.410	35.020	37.566
21	29.615	32.671	36.343	38.932
22	30.813	33.924	37.659	40.289
23	32.007	35.172	38.968	41.638
24	33.196	36.415	40.270	42.980
25	34.382	37.652	41.566	44.314
26	35.563	38.885	42.856	45.642
27	36.741	40.113	44.140	46.963
28	37.916	41.337	45.419	48.278
29	39.087	42.557	46.693	49.588
30	40.256	43.773	47.962	50.892