

การบ้านที่ 2 (1/2564)

รายวิชา คพ.348 แบบจำลองสำหรับวิทยาการข้อมูล

ข้อที่ 1 (ข้อเขียน, 60 คะแนน)

1.1 ให้นักศึกษายกตัวอย่างสถานการณ์ที่มีลักษณะที่เป็นการทดลองสุ่มที่เหมือนกันและเป็นอิสระต่อกัน (Independent and Identical) n ครั้ง มา 3 สถานการณ์ พร้อมอธิบายตามความเข้าใจของนักศึกษาว่าการที่เหตุการณ์แต่ละครั้งเหมือนกันและเป็นอิสระต่อกันหมายความว่าอย่างไร (15 คะแนน)

** ให้ยกตัวอย่างสถานการณ์นอกเหนือจาก เครื่องวัดอุณหภูมิ เครื่องวัดความเร็วแสง การยืนหรือถอย การโยนลูกเต๋า

ការងាររៀងរាល់ទី ២ : ការសម្រេចកូហបងបន្ទាន់ពាណិជ្ជកម្ម និងការសម្រេចកូហបងបន្ទាន់ពាណិជ្ជកម្ម

ចំណាំនៅពេលវេលាដើម្បីនកុងក្រុងប្រជុំសាស្ត្រ ត្រូវបានគ្រប់គ្រង ដើម្បីការអនុវត្តការងារក្នុងក្រុងក្រុងប្រជុំសាស្ត្រ និងក្នុងក្រុងប្រជុំសាស្ត្រ ដែលបានរៀបចំឡើង និងក្នុងក្រុងប្រជុំសាស្ត្រ ដែលបានរៀបចំឡើង

1.2 กำหนดให้ X_i เป็นตัวแปรสุ่มตัวที่ i โดยตัวแปรสุ่มนี้ทั้งหมด n ตัว ให้แสดงวิธีคำนวณ $\bar{X}_n, E[\bar{X}_n], Var(\bar{X}_n)$ (10 คะแนน)

ຈາກ X_i ເປັນຕົວແປຣສຸມຕົວທີ i ໄດ້ມີທຳນົດ ກລັນ ນາຄາ \bar{X}_n ອະດືອກ

$$\bar{X}_n = \frac{X_1 + X_2 + X_3 + \dots + X_n}{n} = \frac{1}{n} (X_1 + X_2 + X_3 + \dots + X_n) = \frac{1}{n} \sum_{i=1}^n X_i$$

ຈາກ X_i ຖະໜາງຕົວ ຈະມີ mean = M ກ່າວໃຫ້ກຽບວ່າ

$$\begin{aligned} E[\bar{X}_n] &= E\left[\frac{1}{n} \sum_{i=1}^n X_i\right] = \frac{1}{n} E\left[\sum_{i=1}^n X_i\right] = \frac{1}{n} \left(\sum_{i=1}^n E[X_i] \right) \\ &= \frac{1}{n} \left(\underbrace{M + M + M + \dots + M}_{n \text{ ຕົວ}} \right) = \frac{nM}{n} = M \end{aligned}$$

mean of \bar{X} ໂດຍກົດ

ຈາກ X_i ຖະໜາງຕົວ ຈະມີ standard deviation = σ ກ່າວໃຫ້ກຽບວ່າ

$$\begin{aligned} Var(\bar{X}_n) &= Var\left(\frac{1}{n} \sum_{i=1}^n X_i\right) = \left(\frac{1}{n}\right)^2 Var\left(\sum_{i=1}^n X_i\right) = \frac{1}{n^2} \left(\sum_{i=1}^n Var(X_i) \right) \\ &= \frac{1}{n^2} \left(\underbrace{\sigma^2 + \sigma^2 + \sigma^2 + \dots + \sigma^2}_{n \text{ ຕົວ}} \right) = \frac{n\sigma^2}{n^2} = \frac{\sigma^2}{n} \end{aligned}$$

Variance of \bar{X} ໂດຍກົດ

ຕົວທີ່ 4 $\bar{X}_n = \frac{X_1 + X_2 + X_3 + \dots + X_n}{n}$

$$E[\bar{X}_n] = M$$

$$Var(\bar{X}_n) = \frac{\sigma^2}{n}$$

1.3 ให้อธิบายสมการต่อไปนี้อย่างละเอียด (20 คะแนน)

$$1.3.1 \lim_{n \rightarrow \infty} P(|\bar{X}_n - \mu| < a) = 1 \text{ (LoLN)}$$

$$\lim_{n \rightarrow \infty} P(|\bar{X}_n - \mu| < a) = 1 \quad \text{នៅ} \quad \lim_{n \rightarrow \infty} P(M-a \leq \bar{X}_n \leq M+a)$$

ន តើ ចំណាំមានការបញ្ជូន

$\lim_{n \rightarrow \infty}$ តើ តាក់ n ផ្លូវកែល 1

$\lim_{n \rightarrow \infty} \dots = 1$ តើ ន ផ្លូវកែល 1 នៅអតាតាកែវ 1

\bar{X}_n តើ តាក់បង្ហាញ ពេលមែន $x_1, x_2, x_3, \dots, x_n$
នៅក្នុងការការកែតាកែវ

n តើ ជា mean ឬ Distribution

$|\bar{X}_n - \mu|$ តើ ជា error កែងការកែវ

a តើ ជា error កំណត់នៅ

$P(|\bar{X}_n - \mu| < a)$ ជា តាកោរនៅក្នុង \bar{X}_n នៅក្នុង $M-a$ និង $M+a$

ជាដែន $\lim_{n \rightarrow \infty} P(|\bar{X}_n - \mu| < a) = 1$

ជា មីនា n ធៀបឱ្យ តាកោរនៅក្នុង \bar{X}_n នៅក្នុង $M-a$ និង $M+a$ ដែលផ្លូវកែល 1

1.3.2 $\lim_{n \rightarrow \infty} F_{Z_n}(a) = \Phi(a)$ (CLT)

$$\lim_{n \rightarrow \infty} F_{Z_n}(a) = \Phi(a)$$

n តើ៖ តាមពេលវេលា

$$\lim_{n \rightarrow \infty} \text{តាម} n \text{ នឹង} \infty$$

$$\lim_{n \rightarrow \infty} \dots = \Phi(a) \text{ តើ } n \text{ នឹង} \infty \text{ និង} \Phi(a)$$

$F(a)$ តើ CDF (Cumulative Distribution Function)

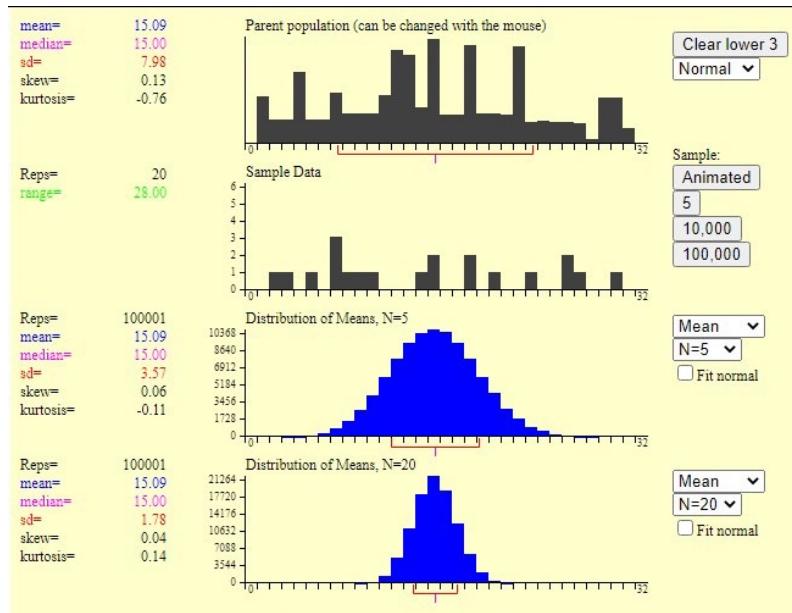
Z_n តើ តាមរបៀបខាងក្រោម Standard normal random variable
តម្លៃ mean (μ) = 0 σ = standard deviation (σ) = 1
 $Z \sim N(0, 1)$ ទីន័រក្នុង $\frac{x-\mu}{\sigma}$

$\Phi(a)$ តើ តាមរបៀប Standard Normal Table

$$\text{តែងតាំង} \lim_{n \rightarrow \infty} F_{Z_n}(a) = \Phi(a)$$

តើ ឱវិជ្ជន ដែលអាចឱ្យបាន តាមរបៀបខាងលើ: ចំណែករាយថាដំឡើង

1.4 ให้นักศึกษาทำการทดลองที่ลิงก์ http://onlinestatbook.com/stat_sim/sampling_dist/ และถ่ายภาพการทดลองของนักศึกษาและอภิปราย



1.4.1 เกี่ยวกับ Central Limit Theorem ของค่า Means ที่พับในการทดลอง (5 คะแนน)

1. การกราฟที่ได้จากการทดลอง

ถ้าตัว n สูงขึ้น ตัวกราฟจะมีลักษณะอย่างนี้ (กราฟด้านบน)

เมื่อ $n \geq 30$ $N\left(\mu, \frac{\sigma^2}{n}\right)$ หมายความว่า การทดลองต่อๆ กัน $\frac{\sigma^2}{n}$

ซึ่งปัจจุบัน ไม่ใช่ตัวกราฟที่เราต้องการ แต่ต้องเป็นแบบนี้

หมายความว่า ตัวกราฟที่ได้มาเราต้อง Population Distribution ลักษณะ

แต่เมื่อทดลองต่อๆ กัน พอ n เป็นจำนวนมาก ตัวกราฟจะเป็นแบบ Normal Distribution

- 1.4.2 เกี่ยวกับความสัมพันธ์ระหว่าง Population distribution กับ Sampling distribution ให้ขัดเจนที่สุดและอธิบายว่า พารามิเตอร์แต่ละค่าที่เกี่ยวข้องกับ Population distribution และ Sampling distribution สำหรับค่า Sample means และ Sample proportions มีอะไรบ้าง รายละเอียดเป็นอย่างไร มีที่มาอย่างไร อย่างละเอียดที่สุด ถ้ามีข้อมูล นอกเหนือจากการทดลองในข้อ 1.3 ต้องอธิบายข้อมูลอื่นนั้นด้วย (10 คะแนน)

ຄວາມສະໝັກພົນຖານ ຈະເປັນ Population distribution ໃລະ Sampling distribution ແລ້ວ

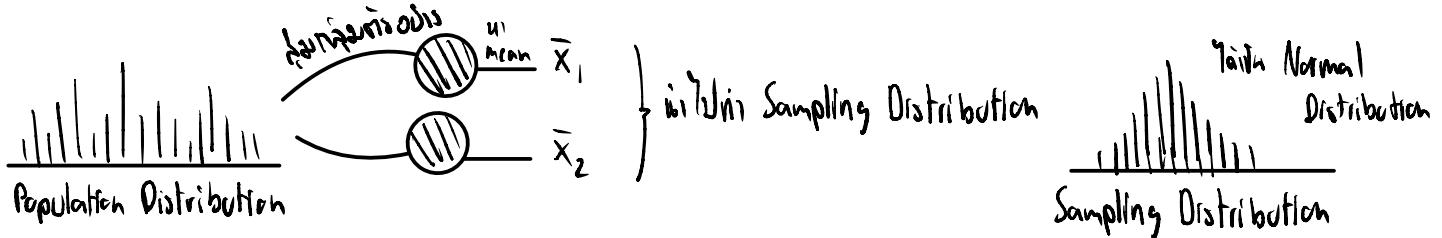
Sampling Distribution ເກືອທິການ ທີ່ມີມາຈຸດຕະວັດປາບຕາກ Population Distribution ແລ້ວນຳໃຫຍ່ mean

ໄລ້ວ່າດີ mean ອາກກົມຕ່ອປິງ ມາ plot graph

ມີນົບຈາກລຸ່ມເຖິງ \bar{X}_1 , ກຳລຸ່ມຈອງເປັນ \bar{X}_2 ໂດຍສືບຕະຫຼາດກຳລຸ່ມລວງໄວ້ໃນ Sampling Distribution

ຕາມໄວ້ວ່າ Population Distribution ດີເລີນອົດຕະກຳ ແກ່ນສັງເກົນມະນຸຍາກີ່ໃນ Sampling Distribution ໄລວ

Sampling Distribution with Normal Distribution នៃទូទាត់ សម្រាប់បង្កើត



ห้องด้วยการเพิ่มปุ่มที่ให้ทั้งการทดลองอุณหภูมิ

Sampling Distribution និងការគូសរួលជាន់ Population Distribution

Tau 1100 = 1000 g; ស្ថិតិ 5 តាម គេងនក់ mean នៃពាន់លម្អិត plot Tauអនុវត្តមាន 10,000 រដ្ឋទីនៅពេលបច្ចុប្បន្ន

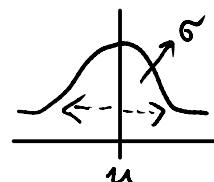
មិនមែនពេលវេលា Sample Means ឬ Sample Proportions ម៉ោង mean ឬ standard deviation តាមទៀតមែនដឹងទៅទៅ តើតុលាន

Sample Means

Parameter	Population Distribution	Sample	Sampling distribution of \bar{X}
Mean	μ	\bar{X}	$\mu_{\bar{X}}, E[\mu_{\bar{X}}] = \mu$
S.D.	σ	s	$\sigma_{\bar{X}} = \frac{\sigma^2}{n}$

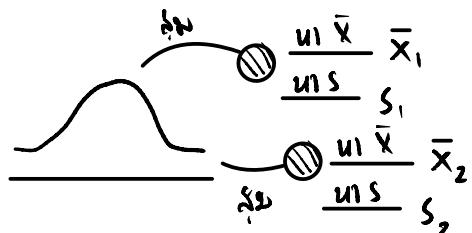
Population Distribution

- តាម mean នឹង μ តើការងារជា mean ទាក់ទងប្រចាំថ្ងៃនៅទៅ ហួនដូរការងារនៅលើ μ
- តាម S.D. នឹង σ តើសម្រេចបន្ថែមនូវភាពរូបបញ្ហាដែលខ្លួន ហួនដូរការងារតូចបែប σ



Sample (តីវការស្តីពីឱ្យលាកកគ្រប់ទាក់ទង)

- តាម mean នឹង \bar{X} ឬអាការារាណជា mean ទៅអីអីប្រចាំថ្ងៃទាក់ទងប្រចាំថ្ងៃ \bar{X}
- តាម S.D. នឹង s ឬអាការារាណសម្រេចបន្ថែមនូវភាពរូបបញ្ហាដែលខ្លួន ហួនដូរការងារតូចបែប s



on CLT

Sampling Distribution of \bar{X} (តីវការ \bar{X} ទាការស្តីពីឱ្យលាកកគ្រប់ទៅ Sampling Distribution) (\bar{X} follow Normal Distribution)

ទៅ Sampling Distribution នាមពេល \bar{X} ទិន្នន័យទៅ CLT រាយការណ៍តូចបែប $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$

ហួនដូរការងារ \bar{X} មិនការកែចាប់បើបាន Normal Distribution តើ mean = μ និង S.D. = $\frac{\sigma}{\sqrt{n}}$

- តាម mean, $\mu_{\bar{X}}, E[\mu_{\bar{X}}]$ នឹង μ តីវការងារ mean ទៅ \bar{X} កំស្ថិត

ទូលាយការរួចរាល់ mean ទៅ Sampling Distribution នឹងតូចបែប mean ទៅការងារ Population Distribution

នៅនេះតាម mean ទៅ Sampling Distribution សម្រាប់ការបែងពាយដែលជាមួយ μ

- តាម S.D., $\sigma_{\bar{X}}$ នឹង $\frac{\sigma}{\sqrt{n}}$ មើលការកែចាប់បើបាន variance ទៅ \bar{X} ចុងក្រោះតើអតិថិជន $\frac{\sigma^2}{n}$

Sample Proportions

Parameter	Population Distribution	Sample	Sampling distribution of \hat{p}
Mean	$M = np$	$\hat{p} = \frac{\bar{x}}{n}$	$E[\hat{p}] = p$
S.D.	$\sigma = \sqrt{npq}$		$\sigma_{\hat{p}} = \sqrt{\frac{pq}{n}}$

ອີງຕະຫຼາດ Population Distribution ສະບັບ Binomial Distribution

ຕີ່ອົງຕະຫຼາດ Binomial ຮຶນຢາກໄດ້ວ່າ ອີ່ mean (M) = np ແລະ S.D. = \sqrt{npq} (ອັນດີຮັບຜົນສົງ Binomial)

ດັ່ງນີ້ Population Distribution

- ອີ່ mean, M ນີ້ $= np$ ພາຍໃຕກົດຮັບຜົນສົງ Binomial (mean = np)
- ອີ່ S.D., σ ນີ້ $= \sqrt{npq}$ ພາຍໃຕກົດຮັບຜົນສົງ Binomial (variance = npq , S.D. = \sqrt{npq})

Sample

- ອີ່ mean, \hat{p} ນີ້ $= \frac{\bar{x}}{n}$ ພາຍໃຕກົດຮັບຜົນສົງ mean ຢູ່ທີ່ມີປະຕິບັດກົດສົມບາ

Sampling Distribution of \hat{p}

- ອີ່ mean ນີ້ $E[\hat{p}]$ ພາຍໃຕກົດຮັບຜົນສົງ mean ອີ່ \hat{p} ກົດສົມບາ
- $E[\hat{p}] = E\left[\frac{\bar{x}}{n}\right] = \frac{1}{n} E[\bar{x}] = \frac{1}{n} E[x] = \frac{np}{n} = p$ E[x] = np ອີ່ Binomial

mean ຍັງ Sampling Distribution ຖືວັດຕ່າງກັນ p

- ອີ່ S.D. ນີ້ $\sigma_{\hat{p}}$ ພາຍໃຕກົດຮັບຜົນສົງ variance ອີ່ \hat{p} ກົດສົມບາ

$$\text{var}(\hat{p}) = \sigma_{\hat{p}}^2 = \text{var}\left(\frac{\bar{x}}{n}\right) = \frac{1}{n^2} (\text{var}(x)) = \frac{1}{n^2} (npq) = \frac{pq}{n}$$

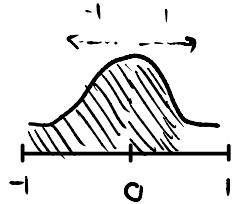
$$\text{ອື່ນດ້ວຍ } \sigma_{\hat{p}} = \sqrt{\frac{pq}{n}}$$

ຂໍ້ທີ 2 (ຂໍ້ເຊີຍ, 40 ຄະແນນ)

2.1 ອີຈິບຍາກເປີດຕາຮັງ Z ແລະ ຕາຮັງ T (ອາຈະຈະວັດຫຼືອຫາດວ່າຍິ່ງເພີ່ມເຕີມຈາກອິນເຕີຣ໌ເນື້ອປະກອບຄໍາອີຈິບຍາ) ວ່າຄ່າທີ່ອຸ່ນໃນ
ຕາຮັງດັກລ່າງໝາຍຖືກຄໍາວ່າໄວ ແລະ ຄ່າທີ່ສອດຄລ້ອງກັບຄ່າທີ່ອຸ່ນໃນຕາຮັງດັກລ່າງມີຄໍາວ່າໄວບ້າງ (5 ຄະແນນ)

ຕາຮັງ Z

Z ມີ Z-score ເປັນຄ່າທີ່ຂອງພື້ນຖານທີ່ໄດ້ການໃຫຍ່ Standard Normal



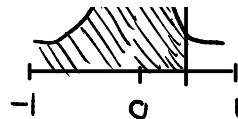
ໄຫຍ້ຕ່າ mean ຂອງ Z ເປັນຕ່າງແນວຂອງຕາມມາແກ້ວ ຖ້າໄດ້ປັນ 0 ເລກຂອງ ແລະ ຕ່າ S.D. ທີ່ໄດ້ປັນ 1 ເລກຂອງ

ດັ່ງນີ້ ຕາຮັງ Z ມີ Z-table ທີ່ເປັນຕາຮັງທີ່ນີ້ພື້ນຖານທີ່ໄດ້ການໃຫຍ່ Standard Normal

ຍົກຄ່ອນຢ່າງເຊັ່ນ ຕ່າ z ພົດຕ່າງກັນ 1.00 ລ່າຍິ່ງໄປເປົ້າໃນຕາຮັງຕ່າ Z

z	.00	.01	.02	.03
0.0	.5000	.5040	.5080	.5120
0.1	.5398	.5438	.5478	.5517
0.2	.5793	.5832	.5871	.5910
0.3	.6179	.6217	.6255	.6293
0.4	.6554	.6591		
0.5	.6915	.6950		
0.6	.7257	.7291		
0.7	.7580	.7611		
0.8	.7881	.7910		
0.9	.8159	.8186		
1.0	<u>.8413</u>	.8438	.8461	.8485
1.1	.8643	.8665	.8686	.8708
1.2	.8849	.8869	.8888	.8907
1.3	.9032	.9049	.9066	.9082
1.4	.9192	.9207	.9222	.9236

ຖືກຕ່າງ ທ່າ z = 1.00 ຖ້າພື້ນຖານທີ່ໄດ້ການ 0.8413

2.2 ອີຈິບຍາກໝາຍຂອງຊ່ວງຄວາມເຂື້ອມໜັ້ນ (Confidence Interval) ວ່າມີຄວາມໝາຍວ່າຍິ່ງໃຫຍ່ ພ້ອມເຂື້ອນສູງຮຽກຮາຫ່າງຄວາມ
ເຂື້ອມໜັ້ນສໍາຫັນ Population Mean ແລະ Population Proportion (5 ຄະແນນ)ຕ່າກ່ານເກີບຂອງກັນຕ່າ Z ສຳເນົາມີ $\mu, \bar{x}, \sigma, s, n$ ດ້ວຍຄໍານີ້ສໍາວັນການໃໝ່ຢູ່ຢັນທີ່ໄດ້ການຕາມ ສັບການຄ່າແນວດັບນີ້ $Z = \frac{\bar{x} - \mu}{\sigma}$ ນີ້ອໍຈຶ່ງແລະ $Z = \text{Confidence Intervals}$ ສໍານັກ Population Mean ສັການໄວ້ກົງນີ້ ພາກທັບ σ , c_s ກ່າວເກີນ mean = $\bar{x} \pm Z^* \frac{\sigma}{\sqrt{n}}$ ແລະ $Z = \text{Confidence Intervals}$ ສໍານັກ $\text{Population Proportion}$ ສັການໄວ້ມັດນີ້ $\hat{p} \pm Z^* \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

ຕາງ T ດິນທາງກໍ່ໄສລາດ ດ້ວຍກູຫາ

ໄລຍກອງຕ່າງມາກາງ T ແນວດຕະຫຼາກໄລຍ Degree of Freedom ໄກສະຕິກຳນົດນີ້ ນີ້ແມ່ນ Significance

$$\text{Degree of Freedom} = n - 1$$

ຕ່າດຈານພ່ອມືນ (Confidence level) ນີ້ແມ່ນ Significance ອົບກັນແລະຂົງກຳນົດ

ສ່ານໃນຢູ່ກົດຕ່າງການ + ຖຸມາຕາມ Confidence Intervals ສຳເນົາ Population Mean ໄກສະຕິ t test

$$C.I \text{ for Mean: } \bar{x} \pm t^* \frac{s}{\sqrt{n}}$$

t test: ເປັນກາງເທິ່ງກົດຕ່າງການ test statistic ພົມກາງ, ດ້ວຍກູຫາກຳນົດ ທີ່Reject
ວິນ test statistic ພົມກາງ, ດ້ວຍກູຫາກຳນົດ ທີ່Retain

ຕ່າງມາກາງຢູ່ຕາງ T

	16	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921
17	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	
18	0.000	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	
19	0.000	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	
20	0.000	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	
21	0.000	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	
22	0.000	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	
23	0.000	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	
24	0.000	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	
25	0.000	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	
26	0.000	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	
27	0.000	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	
28	0.000	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	
29	0.000	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	
30	0.000	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	
40	0.000	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	
60	0.000	0.679	0.848	1.045	1.296	1.671	2.000	2.390	2.660	
80	0.000	0.678	0.846	1.043	1.292	1.664	1.990	2.374	2.639	
100	0.000	0.677	0.845	1.042	1.290	1.660	1.984	2.364	2.626	
1000	0.000	0.675	0.842	1.037	1.282	1.646	1.962	2.330	2.581	
Z	0.000	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	
	0%	50%	60%	70%	80%	90%	95%	98%	99%	
							Confidence Level			

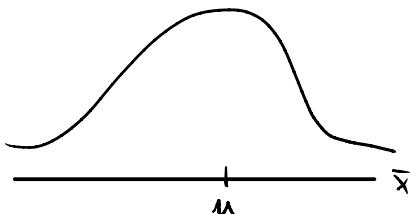
ສະບັບ DF = 20 ||| Confidence level = 95%

ທີ່ຈຳຕ່າງກົດຕ່າງການ 2.086

||| = ດິນທາງກໍ່ໄສລາດ ປັນເປົວຄວາມຫຼັມນີ້ ນີ້ແມ່ນ t test

2.2 ອີບາຍຄວາມໝາຍຂອງຫ່ວງຄວາມເຂື້ອມື້ນ (Confidence Interval) ວ່າມີຄວາມໝາຍວ່າອ່າງໃຈ ພ້ອມເປີຍສູດກາຣາຫ່ວງຄວາມເຂື້ອມື້ນສໍາຫຼັບ Population Mean ແລະ Population Proportion (5 ດະແນນ)

Confidence Interval ຈະຮັດໃຫຍ່ກຳໄປປະຈາກທັງໝົດ ທ່າງອາພາວິເຕີອີກົງປົນດັກວົງຂອງ Population ມີຄວາມ



ນາກກ່າວແນດ C.I ທີ່ 95% ນໍາໃຊ້ກຳຈົດກຳລົມ 95 ຕົວ

ພາວິເຕີອີກົງປົນດັກວົງໃນສ່ວນທີ່ດຳເນັດໄດ້ແລ້ວ 95 ຕົວ

ໄຕນປົກໂຕ ລາຍງົາການ Population Distribution ຮີ່ໄຟ້ຈາກໄມ້ກວາບຕໍ່ mean (μ) ແລະ S.D. (σ)

ແຕ່ເຖິງ 0. ກຽບນີ້ຈະແນ່ໃນຄາຖາກທີ່ໄດ້ observe ຂາ ຫຼືທ່ານ observe ພາກເກີດ Uncertainty ຂາ

ດັ່ງນີ້ ລາຍງົາ C.I ພົບໄຟ້ທີ່ກຳຈົດກຳລົມ ທ່າງອາພາວິເຕີ mean (μ) ແລະ S.D. (σ) ຂອງ Population Distribution ມີຄວາມ

ສູງຕາກການປັບປຸງຄວາມ

2.3 ອີບາຍຄຸນສົມບັດ Random, Normal, Independence ສໍາຫຼັບກາຣີ Inference ສໍາຫຼັບຄໍາ Mean ແລະ Proportion ພອມ
ກຸລຸມປະຫາກ (5 ດະແນນ)

ນາກຮາກາຖານ ດ້ວຍ S.D. ຂອງກຸລຸມປະຫາກ (σ) ຈະໄຟ້ Z test ຜົ່ງຕົ້ນ C.I : $\bar{x} \pm z^* \frac{\sigma}{\sqrt{n}}$

ນາກໄມ້ກວາບຕໍ່ S.D. ວົງກຸລຸມປະຫາກ (σ) ຈະໄຟ້ T test ຜົ່ງຕົ້ນ C.I : $\bar{x} \pm t^* \frac{s}{\sqrt{n}}$ ບໍ່ຈະ $\bar{x} \pm t^* \frac{s}{n}$

ສູງຕາກການປັບປຸງຄວາມເຫັນວ່າສຳຫຼັບ Population Proportion ຈະໄຟ້ Z test ມີຄວາມ C.I : $\hat{p} \pm z^* \sqrt{\frac{\hat{p}\hat{q}}{n}}$

2.3 ခិត្យឱយគុលសមប័តិ Random, Normal, Independence ដោយបង្កើត Inference សំខាន់ការធានាឌាម្បី Mean និង Proportion ទាំងអស់ ៥ គម្រោង

ការធានាឌាម្បី Mean សំខាន់ការធានាឌាម្បី

Random តូចៗ ការកែបានពេលវេលាដែលការបង្ហាញត្រូវបានធ្វើឡើងនៅក្នុងប៊ូតិុលីតិុលី bins

Normal តូចៗ Sampling Distribution ជាតួរប៊ូតិុលី Normal Distribution

(ឬពី Population Distribution របស់វា Normal) និងបានដោះស្រាយ នៅពេល Sample កំណត់តាម plot មួយទូទាត់ការកែបានដោយការបញ្ចូល 30)

Independence តូចៗ ថា mean និងការស្ថិតិយោគ (observation) នឹងត្រូវបានដោះស្រាយ ពេលកែបាន

និងក្នុង Sampling ដែលបានបង្ហាញឡើងតិចៗ Sample size មួយចំនួន 10% ទៅ ៥០% និងក្នុងមេដារ

ការធានាឌាម្បី Proportion សំខាន់ការធានាឌាម្បី

Random តូចៗ ការកែបានពេលវេលាដែលការបង្ហាញត្រូវបានធ្វើឡើងនៅក្នុងប៊ូតិុលី bins

Normal តូចៗ Sampling Distribution ជាតួរប៊ូតិុលី Normal Distribution ($n\hat{p}$ ឬ $n\hat{q}$ នឹងត្រូវបានដោះស្រាយ ពេលកែបាន 10)

ការស្ថិតិយោគ (Normal)

Independence តូចៗ ថា mean និងការស្ថិតិយោគ (observation) នឹងត្រូវបានដោះស្រាយ ពេលកែបាន

និងក្នុង Sampling ដែលបានបង្ហាញឡើងតិចៗ Sample size មួយចំនួន 10% ទៅ ៥០% និងក្នុងមេដារ

2.4 ในการทดลองที่มีข้อเสียงหนึ่งของ Mendel เกี่ยวกับถั่ว คือ เขายืนยันว่าร้อยละ 25 ถั่vrุ่นลูกจะมีสีเหลือง แต่ในสิ่งที่ได้จากการทดลองคือเขาได้ถั่วสีเหลือง 152 และได้สีเขียว 428 ตัว ให้เราระบุอัตราส่วนของถั่วสีเหลืองในรุ่นลูก ให้แสดงวิธีหาช่วงความเชื่อมั่นที่ 95% สำหรับอัตราส่วนของถั่วสีเหลืองในรุ่นลูก ให้แสดงวิธีหาช่วงความเชื่อมั่นดังกล่าวอย่างละเอียด (10 คะแนน)

ตารางบัว Confidence Level = 95%, $Z^* = 1.960$ (จากตารางทาง左)

	0.000	0.005	0.007	0.010	0.015	0.020	0.025	0.030	0.035	0.040	
40	0.000	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	0.000	0.679	0.848	1.045	1.296	1.671	2.000	2.390	2.660	3.232	3.460
80	0.000	0.678	0.846	1.043	1.292	1.664	1.990	2.374	2.639	3.195	3.416
100	0.000	0.677	0.845	1.042	1.290	1.660	1.984	2.364	2.626	3.174	3.390
1000	0.000	0.675	0.842	1.037	1.282	1.646	1.962	2.330	2.581	3.098	3.300
Z	0.000	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	3.090	3.291
	0%	50%	60%	70%	80%	90%	95%	98%	99%	99.8%	99.9%
							Confidence Level				

$$X_1 = 152, X_2 = 428 \quad q = ? \quad n = 152 + 428 = 580$$

โดยมีต้องการหาอัตราส่วนของถั่วสีเหลืองในรุ่นลูก จึงนำ CI คำนวณ Proportion

$$CI : \hat{p} \pm Z^* \sqrt{\frac{\hat{p}\hat{q}}{n}} \rightarrow \text{Var Proportion} = \sqrt{\frac{\hat{p}\hat{q}}{n}}$$

$$\therefore 0.262069 \pm 1.960 \left(\sqrt{\frac{(0.262069)(0.737931)}{580}} \right)$$

$$\therefore [0.226279, 0.297859]$$

$$\hat{p} = \frac{x}{n} = \frac{152}{580} = 0.262069$$

$$\hat{q} = 1 - \hat{p}$$

$$= 1 - 0.262069 = 0.737931$$

ดังนั้น ในช่วงความเชื่อมั่นที่ 95%

อัตราส่วนของถั่วสีเหลืองในรุ่นลูกมีอัตราส่วนอยู่ในช่วง 0.226279 ||A = 0.297859

||A: Mendel ได้ก่อเหตุร้ายไปบุญที่ Mendel ได้ก่อเหตุร้ายไปบุญที่ Mendel

- 2.5 ຜູ້ຈັດຄ່າຍອຮຽມແທ່ງໜຶ່ງທີ່ມີຜູ້ເຂົ້າວ່ານັບພັນຄນສນໃຈວ່າຄ່າເຄີຍຂອງຈຳນວນຈົດໝາຍທີ່ຜູ້ເຂົ້າວ່າເປັນຄື່ນຳຄົງກົງເປັນເທົ່າໄໝ ສມມືວ່າເຮົາຮູ້ວ່າ standard deviation ຂອງຄນໃຈຢ່າງທັງໝົດຄື່ນ 2.5 ຈາກການສອບຄາມຈາກຜູ້ທີ່ອູ້ໃນຄ່າຍ 25 ດັ່ງຕໍ່າງໆ
- 7.9 ຄັບຕ່ອ້ນຶ່ງຄນ ແລະ sample standard deviation ຄື່ນ 2.8 (15 ດັ່ງຕໍ່າງໆ)

2.5.1 ໄທຮບ່າຍຄ່າ \bar{X}, σ, n

$$\bar{X} = 7.9, \sigma = 2.5, n = 25$$

2.5.2 ໄທຮບ່າຍໃນການນິ້ຕັ້ງແປຣສຸມ X ແລະ \bar{X} ມາຍຄື່ນໄວ້

X ດີ່ວ່າ ຈຳນວນ ອອນມາບຖື່ກິ່ງທ່າງມີຕ່າງໆ ເບີນກົດພັງໄນດ້

\bar{X} ດີ່ວ່າ ດ້ວຍລົງຂອງຈຳນວນອອນມາບຖື່ກິ່ງທ່າງມີຕ່າງໆ ເບີນກົດພັງໄນດ້

2.5.3 ໄທສ້າງຂ່າງຄວາມເຂົ້າມື່ນສໍາຫຼັບຄ່າ Population Mean ຂອງຈຳນວນຈົດໝາຍທີ່ຖືກສົງອອກໄປນອກຄ່າຍ

ໜາ C.I ສໍາຫຼັບ Population Mean

ເນື້ອງຈາກກວດ S.D. ຂອງ Population (ມ) ສົ່ງ Z test

ຮູ້ຕ່າງ Z test ດືວ່າ

$$\text{C.I} : \bar{X} \pm Z^* \frac{\sigma}{\sqrt{n}}$$

$$: 7.9 \pm 1.96 \left(\frac{2.5}{\sqrt{5}} \right)$$

$$: [6.92, 8.88]$$

ດ້ວຍກົດລາຍກົດມີຕ່າງໆ 95% ສໍາຫຼັບ Population Mean

ດ້ວຍກົດ 6.92 ໃຊ້ 8.88

ເນື້ອງຈາກໂທນີ້ໄວ້ໄດ້ກົດ Confidence Level

ຕ້ອງໃຫ້ default ເນື້ອງ 95%.

ຕ້ອງ Confidence Level = 95% ເພີ້ນຕໍ່າ = ໃໝ່ຕ່າງໆ

	40	50	60	70	80	90	95%	98%	99%	99.8%	99.9%
40	0.000	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	0.000	0.679	0.848	1.045	1.296	1.671	2.000	2.390	2.660	3.232	3.460
80	0.000	0.678	0.846	1.043	1.292	1.664	1.990	2.374	2.639	3.195	3.416
100	0.000	0.677	0.845	1.042	1.290	1.660	1.984	2.364	2.626	3.174	3.390
1000	0.000	0.675	0.842	1.037	1.282	1.646	1.962	2.330	2.581	3.098	3.300
2000	0.000	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	3.090	3.291
Confidence Level											

ດ້ວຍ Z = 1.96