Test of Normality

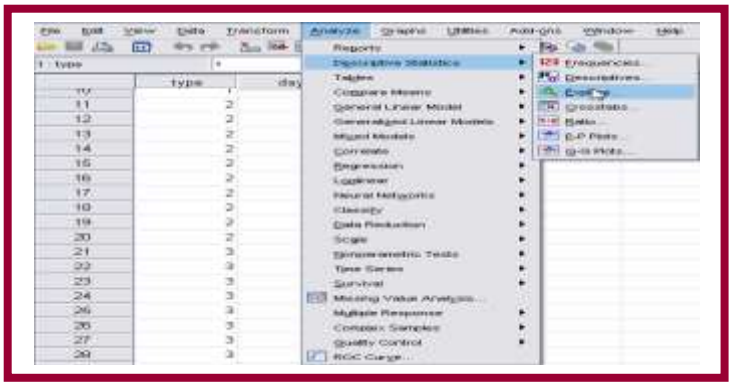
บ่อยครั้งที่ผู้วิจัยดำเนินการทดสอบทางสถิติที่เป็นแบบพาราเมตริก (parametric test)การทดสอบแบบนี้จะมีสมมติฐานกำหนดเอาไว้ว่าประชากรหรือกลุ่มตัวอย่างที่นักวิจัยต้องการศึกษาอยู่นั้นมีการกระจายแบบNormal (Normal distribution) โดยเฉพาะอย่างยิ่งการทดสอบทางสถิติแบบพาราเมตริกที่นักวิจัยส่วนใหญ่คุ้นเคยเช่น t-test , F-test, χ^2 test ล้วนมีสมมติฐานว่าข้อมูลของกลุ่มตัวอย่างที่มาจากประชากรนั้นต้องมีการกระจายแบบ Normal ทั้งนี้เพราะหากข้อมูลมีการกระจายแบบNormal จริง การทดสอบทางสถิติแบบพาราเมตริกจะมีข้อได้เปรียบเหนือกว่าการทดสอบที่ไม่ใช่พาราเมตริกตรงที่ว่าสามารถชี้ชัดให้เห็นความแตกต่างแม้จะใช้ขนาดของกลุ่มตัวอย่างที่เล็กกว่าหรือหากใช้ขนาดของกลุ่มตัวอย่างเท่ากัน การทดสอบแบบพาราเมตริกจะสามารถแยกแยะความแตกต่างแม้จะเล็กน้อยได้ดีกว่าการทดสอบที่ไม่ใช่พาราเมตริก

จะตรวจสอบว่าข้อมูลที่ต้องการศึกษามีการกระจายแบบNormal หรือไม่อย่างไร?

สามารถใช้โปรแกรมสำเร็จรูปทางสถิติ ดังจะแสดงให้เห็นตามตัวอย่างของตารางด้านล่างซึ่งแสดงจำนวนวันกว่าราจะเริ่มเกิดขึ้นกับเนยที่ห่อหุ้มด้วยวัสดุสามชนิด โดยกำหนดว่าขนาดของตัวอย่างเนยที่ห่อหุ้มด้วยวัสดุแต่ละชนิดเท่ากับ 10

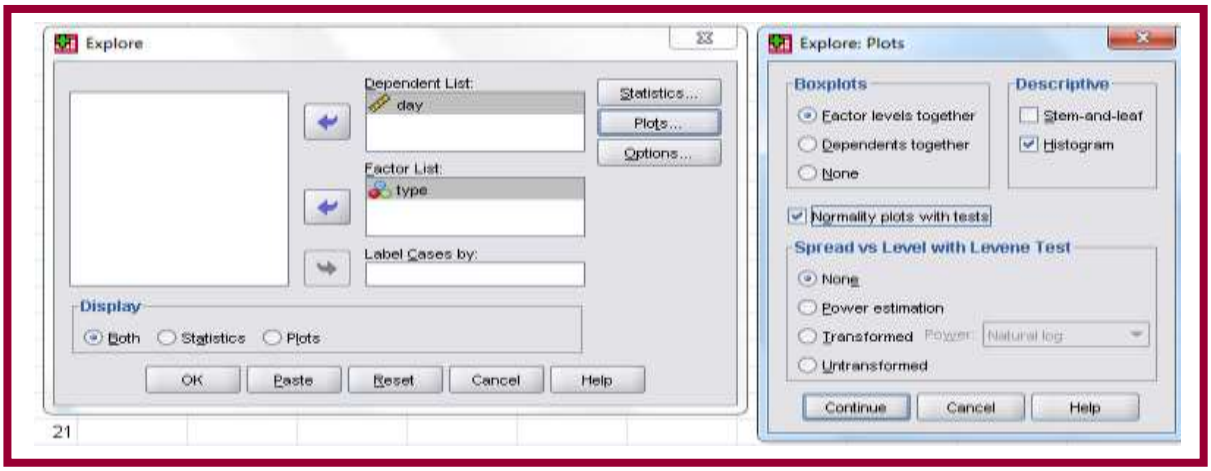
จำนวนวันที่ราจะเริ่มเกิดขึ้นกับเนยที่ห่อหุ้มด้วยวัสดุชนิดต่างๆกัน										
วัสดุชนิดที่ 1	15	23	9	12	25	18	6	21	14	11
วัสดุชนิดที่ 2	10	18	10	8	17	18	4	16	15	5
วัสดุชนิดที่ 3	18	20	25	27	19	20	17	22	21	18

●คลิ๊ก Analyze>Descriptive Statistics >Explore



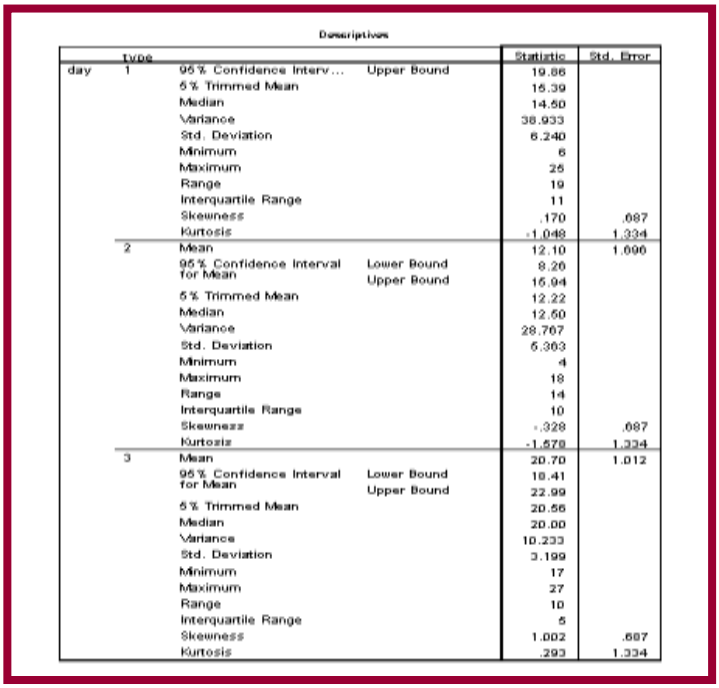
●คลิ๊กday(จำนวนวันกว่าราจะขึ้น) ให้ไปอยู่ในกล่อง Dependent List คลิ๊กtype(ประเภทวัสดุที่ใช้)ให้ไปอยู่ในกล่อง Factor List

•คลิ๊กปุ่ม Plots ในส่วน Descriptive ให้เลือก Histogram แล้วให้เลือกNormality plots with tests



•คลิ๊ก Continue คลิ๊ก OK

Computer output แสดงผลดังนี้

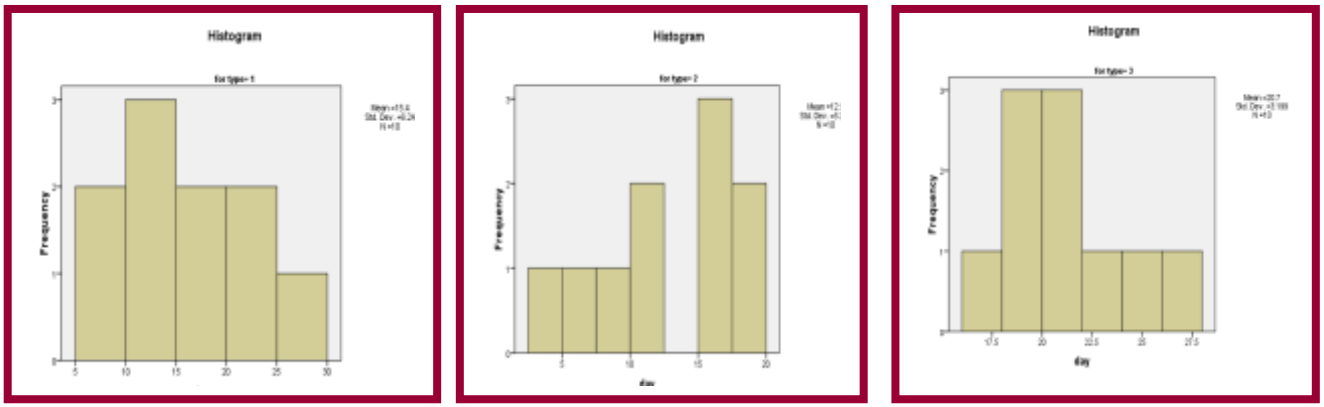


Descriptives

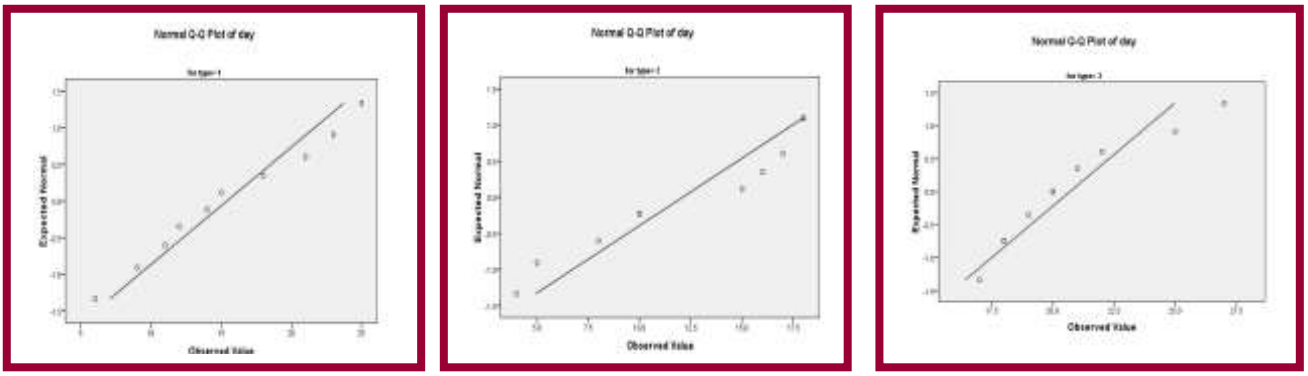
	type			Statistic	Std. Error
day	1	95% Confidence Interv...	Upper Bound	19.86	
		5% Trimmed Mean		16.39	
		Median		14.50	
		Variance		38.933	
		Std. Deviation		6.240	
		Minimum		6	
		Maximum		25	
		Range		19	
		Interquartile Range		11	
		Skewness		.170	.687
		Kurtosis		-1.048	1.334
	2	Mean		12.10	1.096
		95% Confidence Interval for Mean	Lower Bound	8.26	
			Upper Bound	15.94	
		5% Trimmed Mean		12.22	
		Median		12.50	
		Variance		28.767	
		Std. Deviation		5.363	
		Minimum		4	
		Maximum		18	
		Range		14	
		Interquartile Range		10	
		Skewness		-.328	.687
		Kurtosis		-1.579	1.334
	3	Mean		20.70	1.012
		95% Confidence Interval for Mean	Lower Bound	18.41	
			Upper Bound	22.99	
		5% Trimmed Mean		20.56	
		Median		20.00	
		Variance		10.233	
		Std. Deviation		3.199	
		Minimum		17	
		Maximum		27	
		Range		10	
		Interquartile Range		5	
		Skewness		1.002	.687
		Kurtosis		.293	1.334

หากข้อมูลมีการกระจายแบบNormal จริง ความเบ้ (skewness) จะมีค่าอยู่ระหว่าง- 1 ถึง 1 และความโด่ง (kurtosis) จะมีค่าอยู่ระหว่าง -1 ถึง 1 ในที่นี้จำนวนวันกว่าราจะขึ้นของวัสดุชนิดที่หนึ่งและสามจะใกล้เคียงกับการกระจายแบบNormal มากกว่าของวัสดุชนิดที่ 2

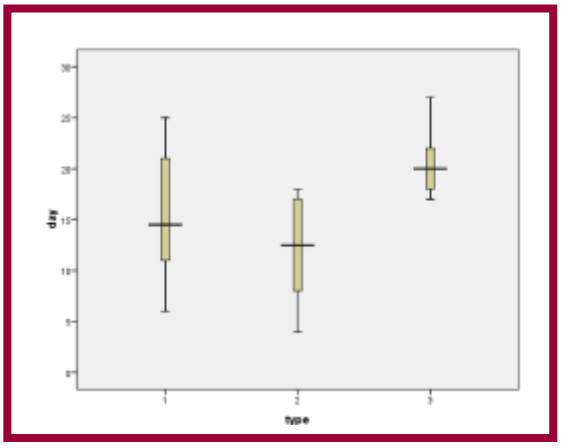
ภาพถัดไปแสดงHistogram แสดงการกระจายของวันกว่าราจะขึ้นของวัสดุทั้งสามชนิด



ภาพถัดไปแสดงQ-Q plotแสดงการกระจายของวันกว่าราจะขึ้นของวัสดุทั้งสามชนิดหากการกระจายเป็นแบบNormal จริงไม่ผิดเพี้ยน จุดต่าง ๆจะอยู่บนแนวเส้นทแยงมุม



ภาพถัดไปแสดงBox-plot แสดงการกระจายของวันกว่าราจะขึ้นของวัสดุทั้งสามชนิด หากการกระจายเป็นแบบNormal จริง เส้นผ่าในแนวนอนซึ่งแสดงค่าเฉลี่ยจะต้องอยู่กึ่งกลางของแท่งเทียน และกึ่งกลางระหว่างปีกกาที่แสดงในแนวตั้ง



ตารางสุดท้ายแสดงผลการทดสอบด้วยค่าสถิติ

Tests of Normality

	type	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Statistic	df	Sig.	Statistic	df	Sig.
day	1	.126	10	.200*	.969	10	.881
	2	.206	10	.200*	.889	10	.167
	3	.187	10	.200*	.908	10	.265

a. Lilliefors Significance Correction

*. This is a lower bound of the true significance.

Kolmogorov-Smirnov(Lilliefors) testที่ใช้สำหรับกลุ่มตัวอย่างขนาดใหญ่และ Shapiro-Wilk test ที่ใช้สำหรับกลุ่มตัวอย่างขนาดเล็ก ต่างยืนยันการทดสอบสมมติฐานที่ว่า การกระจายของข้อมูลไม่ได้แตกต่างจากการกระจายแบบ Normal อย่างมีนัยสำคัญ

Contribution this issue: ดร. ดนัย ปัตตพงศ์

อยากเรียนรู้การนำสถิติข้างต้นนี้ไปใช้ในการวิจัยระดับสารนิพนธ์ (independent study) วิทยานิพนธ์ (thesis) ดุษฎีนิพนธ์(dissertation) ปรึกษาได้ที่ dpattaphongse@gmail.com / dpattaphongse@yahoo.com

* ผู้แต่ง MBA's Made Easy (160+ issues) เอกสารวิชาการด้านศาสตร์การบริหารธุรกิจที่ช่วยให้ธุรกิจสามารถยืนหยัดและอยู่รอดได้ในภาวะที่โลกเปลี่ยนแปลงอยู่ตลอดเวลา

* ผู้พัฒนา FINALYSIS... a dedicated software สำหรับช่วยในการวิเคราะห์ความเป็นไปได้ทางการเงินของโครงการพัฒนาอสังหาริมทรัพย์ (บ้านจัดสรร/จัดสรรที่ดินเพื่อการอุตสาหกรรม/อาคารชุด/อาคารสำนักงานให้เช่า) โรงแรม โรงพยาบาลเอกชน ห้างสรรพสินค้า โรงงานน้ำตาล โรงงานกระดาษ โรงไฟฟ้าชีวมวล ฯลฯ ได้เห็นตัวเลขก่อนโครงการเกิด หลีกเลี่ยงความผิดพลาดเป็นร้อยเป็นพันล้านหากเกิดการลงทุนจริง (กำหนด debut 1 เมษายน 2569)

* ผู้แต่งหนังสือ"การวิเคราะห์ความเป็นไปได้ทางการเงินและการจัดวงเงินเครดิตของโครงการลงทุน"ประกอบด้วยตัวอย่างของธุรกิจจริงที่ไม่เปิดเผยชื่อนับ 100 บริษัท ครอบคลุมอุตสาหกรรม 24 อุตสาหกรรม

*Co-developer ซอฟแวร์ en◉gex◉cel® สำหรับใช้ทดสอบ/เรียนรู้ศัพท์(ประกอบด้วยแบบฝึกหัดและเฉลยกว่า 90 บทครอบคลุมศัพท์ระดับ SAT/IELTS/TOEFL กว่า 12,000 คำ) และไวยากรณ์อังกฤษ (ประกอบด้วยแบบฝึกหัดและเฉลยกว่า 160 บทหรือกว่า 10,000 ข้อครอบคลุมเนื้อหาระดับอุดมศึกษาและTOEFL) มาพร้อมกับไฟล์เสียง/ไฟล์ข้อมูล/ฯลฯ อีกมาก(กำหนด debut 1 เมษายน 2569)