

Chương Trình Giảng Dạy Kinh tế Fulbright

Học kỳ Thu năm 2015

Các Phương Pháp Phân Tích Định Lượng

GỢI Ý ĐÁP ÁN -- BÀI TẬP 3

XÁC SUẤT

Ngày Phát: Thứ ba 20/10/2015

Ngày Nộp: Thứ ba 27/10/2015

Bản in nộp trước 8h20 tại Hộp nộp bài tập trong phòng Lab.

Bản điện tử nộp trước 8h20 tại địa chỉ: <http://www.fetp.edu.vn/vn/tai-nguyen/hoc-vien-hien-tai/>

Câu 1: (25 điểm)

Dịch vụ bán hàng qua điện thoại khá phổ biến trên thị trường hiện nay. Để đánh giá hiệu quả công việc của nhân viên, tổng số cuộc gọi mà các khách hàng đồng ý nghe nhân viên bán hàng giới thiệu về sản phẩm, dịch vụ trong một giờ là tiêu chí được công ty Ad.Ltd sử dụng. Chính sách lương phổ biến cho các nhân viên bán hàng qua điện thoại thường bao gồm 2 khoản: (i) mức lương cứng cố định (số tiền mà nhân viên nhận được không phụ thuộc vào số cuộc gọi mà họ thực hiện thành công) và (ii) lương doanh số (dựa vào số cuộc gọi đạt yêu cầu mà nhân viên thực hiện được). Chính sách chung của công ty Ad.Ltd là nếu số cuộc gọi nằm trong khoảng 1 độ lệch chuẩn so với giá trị kỳ vọng thì nhân viên đó được xem là hoàn thành nhiệm vụ và tính lương doanh số theo đơn giá trung bình công ty đã quy định/cuộc gọi. Nếu số cuộc gọi cao hơn hoặc thấp hơn so với ngưỡng đạt yêu cầu trên, nhân viên sẽ được tính lương doanh số với đơn giá/cuộc gọi cao hơn hoặc thấp hơn so với đơn giá trung bình. Số liệu về số cuộc gọi đạt yêu cầu của ngẫu nhiên 50 nhân viên tại công ty Ad.Ltd như sau:

Cuộc gọi/giờ, (X)	0	1	2	3	4	5	Tổng
Số nhân viên (người)	1	4	10	15	15	5	50
Xác suất, f(X)	2%	8%	20%	30%	30%	10%	100%

Dựa vào thông tin cung cấp ở trên, anh chị hãy thực hiện các yêu cầu cụ thể sau:

- a. Số cuộc gọi mà công ty kỳ vọng ở một nhân viên bán hàng nếu dựa vào kết quả làm việc của các nhân viên trên?

$$E(X) = \mu = \sum x * f(x) = 0 * 2\% + 1 * 8\% + \dots + 5 * 10\% = 3,08 \text{ (cuộc gọi)}$$

- b. Một nhân viên được xem là hoàn thành nhiệm vụ thì số cuộc gọi/giờ phải nằm trong khoảng nào?

$x - \mu$	-3.08	-2.08	-1.08	-0.08	0.92	1.92
$(x - \mu)^2$	9.49	4.33	1.17	0.01	0.85	3.69
$f(x) * (x - \mu)^2$	0.19	0.35	0.23	0.00	0.25	0.37

$$\text{Var}(X) = \sum (x - \mu)^2 * f(x) = 1,39 \text{ (Cuộc gọi}^2\text{)}$$

$$\sigma_x = \sqrt{1,39} = 1,18 \text{ (cuộc gọi)}$$

- Để hoàn thành nhiệm vụ, số cuộc gọi /giờ phải nằm trong khoảng $\mu \pm \sigma_x$, tương đương với khoảng $[1,90 ; 4,26]$ cuộc gọi. Vậy mỗi nhân viên trung bình phải gọi được từ 2 đến 4 cuộc gọi trong 1 giờ làm việc.

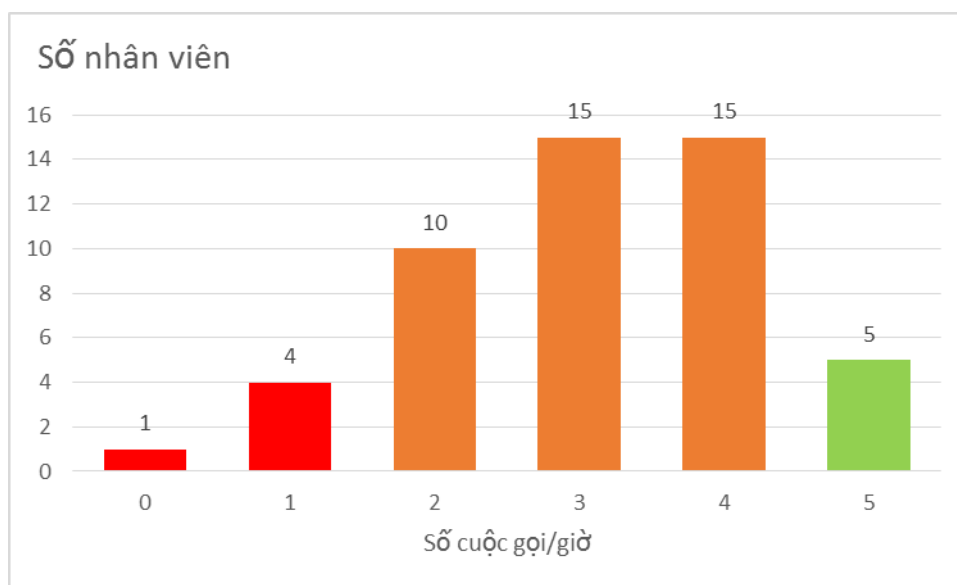
- c. Tính xác suất để một nhân viên được tính lương doanh số trên đơn giá cao hơn đơn giá trung bình?

$$P(X > 4,26) = P(X = 5) = 10\%$$

- d. Tính xác suất để chọn được một nhân viên không thuộc nhóm được đánh giá là hoàn thành nhiệm vụ?

$$P(X < 1,90) = P(X = 0) + P(X = 1) = 2\% + 8\% = 10\%$$

- e. Chọn và vẽ biểu đồ phù hợp cho việc mô tả thông tin tổng hợp ở bảng trên? Nhận xét?



- Nhận xét:

- 80% số nhân viên của công ty hoàn thành nhiệm vụ cho thấy, chính sách của công ty đưa ra hợp lý, vì sẽ không có quá nhiều người bị phạt (ảnh hưởng đến tâm lý của nhân viên khác) và cũng không có quá nhiều người được thưởng (ảnh hưởng đến chi phí của doanh nghiệp, nhưng lại là cơ sở để 80 hay 90% nhân viên còn lại phấn đấu.

Câu 2 (25 điểm):

Điều kiện để thắng cử trong một cuộc bỏ phiếu, ứng cử viên phải có số phiếu ủng hộ đạt từ 65% tổng số phiếu bầu. Trước cuộc tranh cử, ứng cử viên đã tiến hành một cuộc thăm dò ngẫu nhiên với khoảng 500 cử tri trên khắp cả nước. Kết quả thăm dò cho thấy có 280 trong tổng số 500 cử tri được hỏi nói rằng họ sẽ ủng hộ ứng cử viên trên.

- a. Anh/chị hãy tính khoảng xác suất số người ủng hộ ứng cử viên trên với độ tin cậy 95% (hoặc áp dụng quy tắc 2 độ lệch chuẩn)? Khả năng trúng cử của ứng cử viên trên nếu sự ủng hộ của người dân không có thay đổi cho đến ngày tranh cử thực tế?

- Tỷ lệ người ủng hộ trong mẫu thăm dò: $\hat{p} = \frac{280}{500} = 56\%$

- Tỷ lệ người không ủng hộ là: $\hat{q} = 1 - \hat{p} = 1 - 56\% = 44\%$

- Độ lệch chuẩn của phân phối số người ủng hộ là:

$$\sigma = \sqrt{n * p * q} = \sqrt{n * \hat{p} * \hat{q}} = 500 * 56\% * 44\% = 11,1 \text{ (người)}$$

- Số ứng cử viên trong khoảng 2σ là $[280 \pm 2 * 11,1] = [257,8 ; 302,19]$ (người)

- Tỷ lệ ủng hộ ứng cử viên: $p \in \left[\frac{257,8}{500} ; \frac{302,19}{500} \right] \approx [51,56\% ; 60,44\%]$

=> Như vậy, nếu tỷ lệ người ủng hộ không thay đổi, thì khả năng trúng cử của ứng cử viên là không có ($60,44\% < 65\%$ theo yêu cầu).

b. Để hiểu sâu hơn về lý do cử tri ủng hộ hay không ủng hộ cho mình, ứng cử viên cần trao đổi sâu hơn với 5 cử tri được chọn ngẫu nhiên từ 500 cử tri đã tham gia cuộc thăm dò. Tính xác suất để có được ít nhất 3 cử tri trong 5 cử tri được chọn không ủng hộ?

$$P(X \geq 3) = P(X = 3) + P(X = 4) + P(X = 5) = 1 - P(X = 0) - P(X = 1) - P(X = 2)$$

- Ta có:

- $P(X=0) = P(0) = C_0^5 * p^0 * q^{5-0} = \frac{5!}{0! * (5-0)!} * 0.44^0 * 0.56^{(5-0)} = 5,51\%$
- $P(X=1) = P(1) = C_1^5 * p^1 * q^{5-1} = \frac{5!}{1! * (5-1)!} * 0.44^1 * 0.56^{(5-1)} = 21,64\%$
- $P(X=2) = P(2) = C_2^5 * p^2 * q^{5-2} = \frac{5!}{2! * (5-2)!} * 0.44^2 * 0.56^{(5-2)} = 34,00\%$

$$\text{Vậy } P(X \geq 3) = 1 - P(0) - P(1) - P(2) = 38,86\%$$

Câu 3 (30đ)

Tuổi là một trong số các đặc điểm nhân khẩu học đóng vai trò quan trọng, đặc biệt là với những người làm marketing vì hành vi tiêu dùng của mỗi người sẽ khác nhau tùy theo độ tuổi khác nhau (Schwartz, 1990). Một siêu thị vừa khai trương tại thị trấn A với 10 ngàn người đang sinh sống. Số liệu từ cơ quan thống kê tại địa phương cho thấy tuổi trung bình của người dân trong thị trấn là 40 tuổi và độ lệch chuẩn là 15 tuổi. Để xác định lượng hàng hóa nhập về, siêu thị cần phải xác định số lượng người trong từng nhóm tuổi khác nhau. Anh/chị hãy giúp siêu thị A tính toán một số thông tin sau với giả định rằng, tuổi của người dân trong thị trấn tuân theo phân phối chuẩn.

$N = 10.000$ người

$X_{tb} = 40$ (tuổi)

$SD = 15$ (tuổi)

a. Số người dưới 6 tuổi?

$$P(X < 6) = P\left(Z < \frac{6 - 40}{15}\right) = P(Z < -2,26) = \text{normsdist}(-2,26) = 1,17\%$$

$$n = P * N = 1,17\% * 10.000 = 117 \text{ (người)}$$

b. Số người từ 6 đến dưới 18 tuổi?

$$\begin{aligned} P(6 \leq X < 18) &= P\left(\frac{6 - 40}{15} \leq Z < \frac{18 - 40}{15}\right) = P(-2,26 \leq Z < -1,46) \\ &= \text{normsdist}(-1,46) - \text{normsdist}(-2,26) = 5,95\% \end{aligned}$$

$$n = P * N = 5,95\% * 10.000 = 595 \text{ người}$$

c. Số người trên 60 tuổi?

$$P(X > 60) = P\left(Z > \frac{60 - 40}{15}\right) = P(Z > 1,33) = 1 - P(Z < 1,33) \\ = 1 - \text{normsdist}(1,33) = 9,12\%$$

$$n = P \cdot N = 9,12\% \cdot 10.000 = 912 \text{ (người)}$$

d. Tính xác suất số người trong khoảng cộng trừ 1, 2, và 3 độ lệch chuẩn so với tuổi trung bình. Kết quả này có giúp anh chị nhớ tới nội dung nào trong các buổi học trước?

- $P(-1 < Z < 1) = \text{normsdist}(1) - \text{normsdist}(-1) = 68,27\%$
- $P(-2 < Z < 2) = \text{normsdist}(2) - \text{normsdist}(-2) = 95,45\%$
- $P(-3 < Z < 3) = \text{normsdist}(3) - \text{normsdist}(-3) = 99,73\%$

=> Kết quả này tuân theo tính gần đúng của phân phối chuẩn hóa (dạng hình chuông) khi cỡ mẫu lớn.

e. Dữ liệu quá khứ của hệ thống siêu thị cho thấy, khách hàng trong độ tuổi trung niên từ 35 đến dưới 60 tuổi thường là nhóm khách hàng chính. Do vậy, nếu có tối thiểu 50% người dân trong thị trấn thuộc nhóm tuổi này thì việc mở siêu thị tại đây được xem là hợp lý. Anh/chị hãy cho biết, việc mở siêu thị tại thị trấn A trong trường hợp này có hợp lý hay không nếu xét theo tiêu chí trên?

$$P(35 \leq X < 60) = P\left(\frac{35 - 40}{15} \leq Z < \frac{60 - 40}{15}\right) = P(-0,33 \leq Z < 1) \\ = \text{normsdist}(1) - \text{normsdist}(-0,33) = 53,9\% > 50\%$$

Vậy việc mở siêu thị tại thị trấn này là hợp lý.

f. Nếu chỉ lấy đúng 50% số người có tuổi từ 35 trở lên (tuổi được sắp xếp theo thứ tự tăng dần trước khi lấy), người cao nhất trong nhóm này sẽ có độ tuổi là bao nhiêu?

Ta có:

$$P(35 \leq X < X_a) = 50\% \\ \Leftrightarrow P(-0,33 \leq Z < Z_a) = 50\% \\ \Leftrightarrow \text{Normsdist}(Z_a) - \text{Normsdist}(-0,33) = 0,5 \\ \Leftrightarrow \text{Normsdist}(Z_a) = 0,5 + \text{Normsdist}(-0,33) = 0,87 \\ \Leftrightarrow Z_a = 1,12 = \frac{X_a - 40}{15}$$

Vậy $X_a = 1,12 \cdot 15 + 40 = 56,85 \approx 57$ (tuổi) hay $P(35 < X < 57) = 50\%$ (tổng số người)

Câu 4 (20đ)

Thông tin khảo sát một số học viên MPP8 về thời gian làm bài tập của 2 môn học Định lượng và Chính sách Công như sau:

<i>ĐVT: giờ</i>	TG suy nghĩ		TG trình bày	
Môn học	Trung bình	Độ lệch chuẩn	Trung bình	Độ lệch chuẩn
Định lượng (ĐL)	6,0	1,0	4,0	2,0
Chính sách công (CSC)	8,0	2,0	1,0	0,2

Thời gian suy nghĩ và thời gian làm bài là độc lập với nhau. Hệ số tương quan giữa thời gian hoàn thành bài tập Định lượng và thời gian làm bài Chính sách công là 0,7.

Thông tin đề bài được tổng hợp như bảng sau:

<i>ĐVT: giờ</i>	TG làm suy nghĩ (T)		TG trình bày (P)	
Môn học	Trung bình	Độ lệch chuẩn	Trung bình	Độ lệch chuẩn
Định lượng (ĐL)	6	1	4	2
Chính sách công (CSC)	8	2	1	0.2
$r_{DL \cap CSC} = 0.7$				
Thời gian suy nghĩ và trình bày của mỗi môn là độc lập				

Thời gian làm bài môn Định lượng	Kết quả	ĐVT
$E(ĐL) = E(T+P) = E(T) + E(P) =$	10.0	giờ
$Var(ĐL) = Var(T+P) = Var(T) + Var(P) =$	5.0	giờ ²
$SD(ĐL) =$	2.2	giờ

Thời gian làm bài môn Chính sách Công	Kết quả	ĐVT
$E(CSC) = E(T+P) = E(T) + E(P) =$	9.0	giờ
$Var(CSC) = Var(T+P) = Var(T) + Var(P) =$	4.0	giờ ²
$SD(CSC) =$	2.0	giờ

Tổng thời gian làm bài của 2 môn	Kết quả	ĐVT
$E(2 \text{ môn}) = E(ĐL) + E(CSC) =$	19.0	giờ
$Var(2 \text{ môn}) = Var(ĐL) + Var(CSC) + 2Cov(ĐL, CSC) =$	15.33	giờ ²
Trong đó: $Cov(ĐL, CSC) = r_{DL \cap CSC} * SD(ĐL) * SD(CSC)$		
$SD(2 \text{ môn}) =$	3.92	giờ

Dựa vào thông tin cung cấp ở trên, anh/chị hãy tính toán các yêu cầu cụ thể sau đây:

- Xác suất để anh/chị dành không quá 24 giờ để hoàn thành cả 2 bài tập trên (giả định rằng, trong lúc làm bài tập môn ĐL thì anh/chị không thể suy nghĩ hay làm bài tập môn CSC và ngược lại)?

$$P(X < 24) = P\left(Z < \frac{24 - 19}{3.92}\right) = P(Z < 1.28) = \text{normsdist}(1.28) = 89.9\%$$

b. Xác suất để thời gian hoàn thành bài môn ĐL nhiều hơn môn CSC là bao nhiêu?

Gọi D là biến cố chênh lệch thời gian làm bài giữa 2 môn		
$E(D) = E(\text{ĐL}) - E(\text{CSC}) =$	1.0	giờ
$\text{Var}(D) = \text{Var}(\text{ĐL}) + \text{Var}(\text{CSC}) - 2\text{Cov}(\text{ĐL}, \text{CSC}) =$	2.748	giờ ²
$\text{SD}(D) =$	1.66	giờ

$$P(D > 0) = P\left(Z > \frac{0 - 1}{1.66}\right) = P(Z > -0.6) = 1 - \text{Normsdist}(-0.6) = 72.7\%$$

---HẾT---