COMPUTER PROJECT

XÁC SUẤT THỐNG KÊ

Đề tài: Phân tích về thời gian chơi và xếp hạng của người chơi

Lớp: SE1606

Giảng viên hướng dẫn: Nguyễn Việt Anh

Nhóm sinh viên:

HE153639 – Nguyễn Quang Huy

HE153629 – Nguyễn Hà Quốc Duy

HE153561 – Mai Hoàng Anh

Mục lục

I.Xác định đề tài

a, Giới thiệu

- *Tựa game Liên Minh Huyền Thoại(LMHT) là một trong những tựa game nổi tiếng nhất hiện nay thu hút nhiều game thủ*

b, Lí do

- *Đề tài mà nhóm quan tâm là nếu bạn chơi game nhiều thì thứ hạng của bạn có cao hơn không hay cần một lý do khác như trí thông minh hay kĩ năng của bạn tác động.*
- *Sau khi lấy dữ liệu từ 19 bạn ngẫu nhiên có chơi cùng một game*

II, Dữ liệu

- Dữ liệu của nhóm được lấy trực tiếp từ 19 bạn với những câu hỏi được đặt ra để thu được dữ liệu về thời gian chơi game cũng như thứ hạng (ranking) của các bạn

- Thời gian chơi 1 ngày:

Thời gian chơi 1 ngày(giờ)	số lượng người
1	3
2	2
3	3
4	7
5	3
7	1

- Xếp hạng của người chơi:

Ranking	Số lượng người
Sắt (1 điểm)	3
Đồng(2 điểm)	0
Bạc (3 điểm)	2
Vàng (4 điểm)	9
Bạch Kim (5 điểm)	3
Kim Cương (6 điểm)	2

Xử lý dữ liệu:

Parameter	Trung bình	Độ lệch chuẩn	Min	Max
Thời gian chơi (giờ)	3,47	1,59	1	7
Ranking	3,79	1,47	1	6

III, Thống kê mô tả

1, Tìm khoảng tin cậy số giờ chơi và mức rank đạt được của sinh viên

- *Khoảng tin cậy 2 phía μ cho việc thời gian chơi game: (α =0,05):*

Áp dụng công thức:

$$\bar{x} - t_{\alpha/2,n-1}s/\sqrt{n} \leq \mu \leq \bar{x} + t_{\alpha/2,n-1}s/\sqrt{n}$$

Ta có:

Khoảng tin cậy 2 μ cho việc thời gian chơi game:
x=3,47, α = *0,05,* s=1,59 , n=19, $T\alpha/2,19$ = 2,101
Thì khoảng tin cậy là: 2,703 ≤ *μ ≤ 4,236*

- *Khoảng tin cậy 2 phía μ cho xếp hạng của người được khảo sát: (α =0,05):*

Áp dụng công thức:

$$\bar{x} - t_{\alpha/2,n-1}s/\sqrt{n} \leq \mu \leq \bar{x} + t_{\alpha/2,n-1}s/\sqrt{n}$$

Ta có:
Khoảng tin cậy 2 μ cho xếp hạng của người được khảo sát:
x=3,79, α = *0,05,* s=1,47,n=19, $T\alpha/2,19$ = 2,101

Thì khoảng tin cậy là: 3,081≤ μ ≤ 4,498

- Ước lượng 6^2

Áp dụng công thức:

$$\frac{(n-1)s^2}{\chi^2_{\alpha/2,n-1}} \le \sigma^2 \le \frac{(n-1)s^2}{\chi^2_{1-\alpha/2,n-1}}$$

Ta có:

n= 19, s = 1,47 , $X^2{}_{\alpha/2,18}$ = 31,53; $X^2{}_{1-\alpha/2,18}$ =8,23 (Tra bảng chi-square)

=>1,23 ≤ 6^2 ≤4,72

2, Khoảng tin cậy số giờ chơi và rank đặt được của sinh viên nam và nữ

Case 1: $\boldsymbol{\sigma_1}^2$ = $\boldsymbol{\sigma_2}^2$ = $\boldsymbol{\sigma_0}^2$

If $\bar{x}_1, \bar{x}_2, s_1^2$ and s_2^2 are the sample means and variances of two random samples of sizes n_1 and n_2, respectively, from two independent normal populations with unknown but equal variances, then a **100(1 − α)% confidence interval on the difference in means $\mu_1 - \mu_2$ is**

$$\bar{x}_1 - \bar{x}_2 - t_{\alpha/2,n_1+n_2-2}\, s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

$$\le \mu_1 - \mu_2 \le \bar{x}_1 - \bar{x}_2 + t_{\alpha/2,n_1+n_2-2}\, s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} \qquad (10\text{-}19)$$

where $s_p = \sqrt{[(n_1-1)s_1^2 + (n_2-1)s_2^2]/(n_1+n_2-2)}$ is the pooled estimate of the common population standard deviation, and $t_{\alpha/2,n_1+n_2-2}$ is the upper $\alpha/2$ percentage point of the t distribution with $n_1 + n_2 - 2$ degrees of freedom.

A, Khoảng tin cậy 2 phía về sự khác nhau giữa thời gian chơi trung bình của nam và nữ

Với $\overline{x_1}$ = 3.85, n_1 = 13, s_1 = 1.34 (nam)
$\overline{x_2}$ = 2.67, n_2 = 6, s_2 = 1.86 (nữ)
$t_{0.025,17}$= 2,11

Ước tính tổng hợp $s_p = 1.51$

Khoảng tin cậy -0.4 $\leq \mu_1 - \mu_2 \leq$ 2.75

B, Khoảng tin cậy 2 phía về sự khác nhau giữa Rank trung bình của nam và nữ

Với $\overline{x_1}$ = 3.69, n_1 = 13, s_1 = 1.75 (nam)
$\overline{x_2}$ = 4 , n_2 = 6, s_2 = 0.63 (nữ)
$t_{0.025,17}$= 2,11

Ước tính tổng hợp $s_p = 1.5$

Khoảng tin cậy -1.87 $\leq \mu_1 - \mu_2 \leq$ 1.52

Case 2: $\sigma_1{}^2 \neq \sigma_2{}^2$

If $\overline{x}_1$, $\overline{x}_2$, s_1^2, and s_2^2 are the means and variances of two random samples of sizes n_1 and n_2, respectively, from two independent normal populations with unknown and unequal variances, an approximate $100(1-\alpha)\%$ confidence interval on the difference in means $\mu_1 - \mu_2$ is

$$\overline{x}_1 - \overline{x}_2 - t_{\alpha/2,v}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \leq \mu_1 - \mu_2 \leq \overline{x}_1 - \overline{x}_2 + t_{\alpha/2,v}\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} \quad (10\text{-}20)$$

where v is given by Equation 10-16 and $t_{\alpha/2,v}$ is the upper $\alpha/2$ percentage point of the t distribution with v degrees of freedom.

$$v = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(s_1^2/n_1\right)^2}{n_1 - 1} + \frac{\left(s_2^2/n_2\right)^2}{n_2 - 1}}$$

A, Khoảng tin cậy 2 phía về sự khác nhau giữa thời gian chơi trung bình của nam và nữ

Bậc tự do v = 8

Với $\overline{x_1}$ = 3.85, n_1 = 13, s_1 = 1.34 (nam)

$\overline{x_2}$ = 2.67, n_2 = 6, s_2 = 1.86 (nữ)

$t_{0.025,8}$ = 2.306

Khoảng tin cậy $-0.77 \le \mu_1 - \mu_2 \le 3.13$

B, Khoảng tin cậy 2 phía về sự khác nhau giữa Rank trung bình của nam và nữ

Với $\overline{x_1}$ = 3.69, n_1 = 13, s_1 = 1.75 (nam)

$\overline{x_2}$ = 4 , n_2 = 6, s_2 = 0.63 (nữ)

$t_{0.025,8}$ = 2.306

Khoảng tin cậy $-1.58 \le \mu_1 - \mu_2 \le 0.95$

IV, Kiểm định giả thuyết

a, Tin tưởng 95% thời gian chơi trung bình của sinh viên cả trường là 3 giờ/ngày.

1, Tham số quan tâm là hệ số thay thế trung bình là μ

2, H0: µ = *3* giờ/ngày

3, H1: *µ ≠ 3 giờ/ngày*

4, α=0.05

5, Thống kê kiểm định

$$t_0 = \frac{\overline{x} - \mu_0}{s/\sqrt{n}}$$

6, Phản đối H0 if t_0>2.101 hoặc t_0<-2.101

7, Từ $\overline{\boldsymbol{x}}$ = 3,47,s= , $\boldsymbol{\mu_0}$ = 3 and n=19

$t_0 = \frac{3,47-3}{1.59/\sqrt{19}}$ = 1.288

8 Vì t_0 < 2.101 nên không thể bác bỏ H0

b, Tin tưởng 95% Rank trung binh của sinh viên cả trường là 3.

1, Tham số quan tâm là hệ số thay thế trung bình là μ

2, H0: µ = *3*

3, H1: *µ ≠ 3*

4, α=0.05

5, Thống kê kiểm định

$$t_0 = \frac{\overline{x} - \mu_0}{s/\sqrt{n}}$$

6, Phản đối H0 if t_0>2.101 hoặc t_0<-2.101

7, Từ $\overline{\boldsymbol{x}}$ = 3,79,s= , $\boldsymbol{\mu_0}$ = 3 and n=19

$t_0 = \frac{3,79-3}{1.47/\sqrt{19}}$= 2.34

8 Vì t_0 > 2.101 nên bác bỏ H0

C, Kiểm định thời gian chơi trung bình của nam và nữ có bằng nhau không

Với $\overline{\boldsymbol{x_1}}$ = 3.85, $\boldsymbol{n_1}$ = 13, $\boldsymbol{s_1}$ = 1.34 (nam)

$\overline{\boldsymbol{x_2}}$ = 2.67, $\boldsymbol{n_2}$ = 6, $\boldsymbol{s_2}$ = 1.86 (nữ)

$\boldsymbol{t_{0.025,17}}$= 2,11

Case 1: $\boldsymbol{\sigma_1}^2 = \boldsymbol{\sigma_2}^2 = \boldsymbol{\sigma_0}^2$

1, H_0: μ_1 - μ_2 = 0

2, H_1: μ_1 - μ_2 $\neq$ 0

3, α = 0.05

4, *Thống kê kiểm định*

$$t_0 = \frac{\overline{x}_1 - \overline{x}_2 - 0}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

5, Bác bỏ H_0 nếu $t_0 > t_{0.025,17}$ = 2.11 hoặc nếu $t_0 < -t_{0.025,17}$ = -2.11

6,

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

S_p = 1.51 , t_0 = 1.58

7,

Vì $-t_{0.025,17}$ $t_0 < t_{0.025,17}$ nên không thể bác bỏ H_0

Case 2: $\sigma_1{}^2 \neq \sigma_2{}^2$

1, H_0: $\mu_1 - \mu_2 = 0$

2, H_1: $\mu_1 - \mu_2 \neq 0$

3, α = 0.05

4. Thống kê kiểm định

$$T_0^* = \frac{\overline{X}_1 - \overline{X}_2 - \Delta_0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

5, Bậc tự do v = 8 (Đã tính bên trên)

Bác bỏ H_0 nếu $t_0 > t_{0.025,8}$ = 2.306 hoặc nếu $t_0 < -t_{0.025,8}$ = - 2.306

6, t_0 = 1.395

7, **Vì $-t_{0.025,8} < t_0 < t_{0.025,8}$ nên không thể bác bỏ H_0**

D, Kiểm định Rank trung bình của nam và nữ có bằng nhau không

Với $\overline{x_1}$ **= 3.69,** n_1 **= 13,** s_1 **= 1.75 (nam)**

$\overline{x_2}$ **= 4 ,** n_2 **= 6,** s_2 **= 0.63 (nữ)**

$t_{0.025,17}$ **= 2,11**

Case 1: $\sigma_1{}^2 = \sigma_2{}^2 = \sigma_0{}^2$

1, H_0**:** $\mu_1 - \mu_2 = 0$

2, H_1**:** $\mu_1 - \mu_2 \neq 0$

3, α **= 0.05**

4, *Thống kê kiểm định*

$$t_0 = \frac{\overline{x}_1 - \overline{x}_2 - 0}{s_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

5, Bác bỏ H_0 **nếu** $t_0 > t_{0.025,17}$ **= 2.11 hoặc nếu** $t_0 < -t_{0.025,17}$ **= -2.11**

6,

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

S_p = 1.51 , t_0 = - 0.41

7,

Vì - $t_{0.025,17}$ $t_0 <$ $t_{0.025,17}$ **nên không thể bác bỏ** H_0

Case 2: $\boldsymbol{\sigma_1}^2 \neq \boldsymbol{\sigma_2}^2$

1, $\boldsymbol{H_0}$**:** $\boldsymbol{\mu_1}$ **-** $\boldsymbol{\mu_2}$ **= 0**

2, $\boldsymbol{H_1}$**:** $\boldsymbol{\mu_1}$ **-** $\boldsymbol{\mu_2}$ $\neq$ **0**

3, $\boldsymbol{\alpha}$ **= 0.05**

4. Thống kê kiểm định

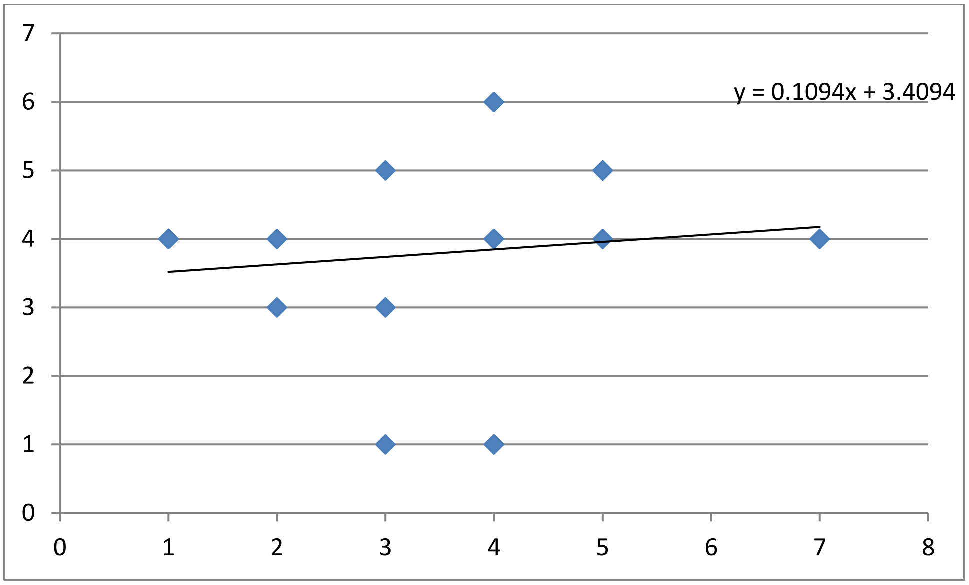
$$T_0^* = \frac{\overline{X}_1 - \overline{X}_2 - \Delta_0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

5, Bậc tự do v = 8 (Đã tính bên trên)

Bác bỏ $\boldsymbol{H_0}$ **nếu** $t_0 > \boldsymbol{t_{0.025,8}}$ **= 2.306 hoặc nếu** $t_0 < -\boldsymbol{t_{0.025,8}}$ **= - 2.306**

6, t_0 = -0.56

7, **Vì** **-** $\boldsymbol{t_{0.025,8}}$ **<** $t_0 <$ $\boldsymbol{t_{0.025,8}}$ **nên không thể bác bỏ** $\boldsymbol{H_0}$



Coefficients	*Standard Error*
3.409411765	0.85577722
0.109411765	0.225352428

Kiểm định giả thuyết

H_0 : β_1= 0

H_1 : $\beta_1 \neq 0$

Với α = 0.05 , $\overline{\beta_1}$ = 0.1094 , se($\overline{\beta_1}$) = 0.225

$$t_0 = \frac{\hat{\beta}_1}{\sqrt{\hat{\sigma}^2/S_{xx}}} = \frac{\hat{\beta}_1}{se(\hat{\beta}_1)}$$

t_0 = 0.486

Với $\boldsymbol{t_{0.025,17}}$ = **2.11 thì** t_0 vẫn nằm trong khoảng -2.11, 2.11

Không thể bác bỏ H_0

Kết luận: Khi mà bạn chơi nhiều một trò chơi thì xếp hạng của bạn sẽ cao hơn