Natalie Julian

**Ejercicio 1: Anova Two way aleatorio**
En una empresa de ensamblaje de metales resulta problemática la variabilidad en la precisión obtenida en cada fusión de partes metálicas. En cada procedimiento, el ideal es ensamblar dos piezas de manera perfecta, de modo de obtener un milimetraje sobrante cero, es decir, que ambas piezas se ensamblen de manera perfecta. Sin embargo, esto no ocurre y siempre se observa un margen de error en cada procedimiento. Se cree además que existe variabilidad en la precisión por máquina de ensamblaje y estación de ensamblaje, para lo cual se seleccionan aleatoriamente 3 estaciones de prueba y 3 máquinas de emsamblaje de la empresa y se midió el sobrante posterior a cada emsamble (error de ensamblaje en centímetros). El experimento se repitió 3 veces. La información se resume en la siguiente tabla:

| **Factor** | **Testing Station** | | |
|---|---|---|---|
| **Assembly Machine** | **B1** | **B2** | **B3** |
| **A1** | 2.3 | 3.7 | 3.1 |
| | 3.4 | 2.8 | 3.2 |
| | 3.5 | 3.7 | 3.5 |
| **A2** | 3.5 | 3.9 | 3.3 |
| | 2.6 | 3.9 | 3.4 |
| | 3.6 | 3.4 | 3.5 |
| **A3** | 2.4 | 3.5 | 2.6 |
| | 2.7 | 3.2 | 2.6 |
| | 2.8 | 3.5 | 2.5 |

a) ¿Qué modelo le propone a la empresa? Justifique su respuesta

Respuesta

Note que los dos factores involucrados son estaciones de prueba y máquinas de ensamblajes, ambos factores son aleatorios pues se eligen aleatoriamente de entre todas las opciones que poseía la empresa. Además, la variable respuesta corresponde al sobrante de cada ensamble, es decir, a la no precisión o error de ensamblaje en el procedimiento. Los experimentos se repitieron 3 veces. Por ende, el modelo que se propone es el siguiente:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk}, \text{con} \quad i = 1, ..., a = 3, \quad j = 1, ..., b = 3 \;\; \text{y} \;\; k = 1, ..., 3$$

$$\alpha_i \overset{\text{i.i.d}}{\sim} N(0, \sigma^2_\alpha)$$

$$\beta_j \overset{\text{i.i.d}}{\sim} N(0, \sigma^2_\beta)$$

$$(\alpha\beta)_{ij} \overset{\text{i.i.d}}{\sim} N(0, \sigma^2_{\alpha\beta})$$

$$\epsilon_{ijk} \overset{\text{i.i.d}}{\sim} N(0, \sigma^2)$$

b) Realice el ajuste del modelo. ¿Es adecuado el modelo propuesto? Justifique su respuesta. En caso de no ser adecuado el modelo, obtenga el modelo adecuado para el análisis del problema

Respuesta

Se definen los datos:

```
(error<-c(2.3, 3.4, 3.5, 3.5, 2.6, 3.6, 2.4, 2.7, 2.8, 3.7, 2.8, 3.7, 3.9, 3.9, 3.4, 3.5,
3.2, 3.5, 3.1, 3.2, 3.5, 3.3, 3.4, 3.5, 2.6, 2.6, 2.5))
 [1] 2.3 3.4 3.5 3.5 2.6 3.6 2.4 2.7 2.8 3.7 2.8 3.7 3.9 3.9 3.4 3.5 3.2 3.5 3.1
[20] 3.2 3.5 3.3 3.4 3.5 2.6 2.6 2.5

(machine<-factor(rep(c(1,2,3), each=3, length=27)))
 [1] 1 1 1 2 2 2 3 3 3 1 1 1 2 2 2 3 3 3 1 1 1 2 2 2 3 3 3
Levels: 1 2 3

(estacion<-factor(rep(c(1,2,3),each=9)))
 [1] 1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 3 3 3 3 3 3 3 3 3
Levels: 1 2 3

model<-aov(error~machine*estacion)

anova(model)[,1:3]
                 Df  Sum Sq  Mean Sq
machine           2 1.60222 0.80111
estacion          2 1.44667 0.72333
machine:estacion  4 0.39778 0.09944
Residuals        18 2.46000 0.13667
```

Se tienen los siguientes tests de interés:
$H_0 : \sigma^2_\alpha = 0,\ H_1 : \sigma^2_\alpha > 0$:

$$F_\alpha = \frac{MCA}{MCAB} \sim F^{1-\alpha}_{(a-1),\ (a-1)(b-1)}$$

```
(Fmachine<-anova(model)[1,3]/anova(model)[3,3])
[1] 8.055866

(a<-length(levels(machine)))
[1] 3

(b<-length(levels(estacion)))
[1] 3

(n<-unique(table(machine,estacion)))  #Caso balanceado
[1] 3

alpha=0.05
Fmachine>qf(p=(1-alpha),df1=(a-1),df2=(a-1)*(b-1))
[1] TRUE
```

Se rechaza la hipótesis de variabilidad nula en las máquinas aleatoriamente seleccionadas.

$H_0 : \sigma^2_\beta = 0,\ H_1 : \sigma^2_\beta > 0$:

$$F_\beta = \frac{MCB}{MCAB} \sim F^{1-\alpha}_{(b-1),\ (a-1)(b-1)}$$

```
(Festacion=anova(model)[2,3]/anova(model)[3,3])
[1] 7.273743

Festacion>qf(p=(1-alpha),df1=(b-1),df2=(a-1)*(b-1))
[1] TRUE
```

Se rechaza la hipótesis de variabilidad nula en las estaciones de prueba aleatoriamente seleccionadas.

Y por último, la interacción $H_0 : \sigma^2_{\alpha\beta} = 0$, $H_1 : \sigma^2_{\alpha\beta} > 0$:

$$F_{\alpha\beta} = \frac{MCAB}{MCE} \sim F^{1-\alpha}_{(a-1)(b-1),ab(n-1)}$$

```
(Finter=anova(model)[3,3]/anova(model)[4,3])
[1] 0.7276423

Finter>qf(p=(1-alpha),df1=(a-1)*(b-1),df2=a*b*(n-1))
[1] FALSE
```

No se rechaza la variabilidad nula asociada a combinación estación y máquina aleatoriamente seleccionadas. Se debe considerar un modelo aditivo.

c) En base al modelo anterior obtenga estimaciones puntuales de las variabilidades involucradas. Comente e interprete las cifras obtenidas.

Respuesta

Dado que se probó que la interacción no resultaba aportar variabilidad al modelo utilizando el test F, se propone el modelo aditivo.

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \epsilon_{ijk}, \text{con} \quad i = 1, ..., a = 3, \quad j = 1, ..., b = 3 \;\; \text{y} \;\; k = 1, ..., 3$$

$$\alpha_i \overset{\text{i.i.d}}{\sim} N(0, \sigma^2_\alpha)$$

$$\beta_j \overset{\text{i.i.d}}{\sim} N(0, \sigma^2_\beta)$$

$$\epsilon_{ijk} \overset{\text{i.i.d}}{\sim} N(0, \sigma^2)$$

Por lo tanto, existen tres fuentes de variabilidad en el modelo propuesto: $\sigma^2$, $\sigma^2_\alpha$ y $\sigma^2_\beta$, observando la tabla anova se pueden estimar puntualmente:

```
aditivo<-aov(error~machine+estacion)
anova(aditivo)[,1:3]
          Df Sum Sq Mean Sq
machine    2 1.6022 0.80111
estacion   2 1.4467 0.72333
Residuals 22 2.8578 0.12990
```

$$\hat{\sigma^2} = MCE = 0.1299$$

$$\hat{\sigma}^2_\alpha = \frac{MCA - MCE}{nb} = \frac{0.80111 - 0.1299}{3 \cdot 3} = 0.07457889$$

$$\hat{\sigma}^2_\beta = \frac{MCB - MCE}{na} = \frac{0.72333 - 0.1299}{3 \cdot 3} = 0.06593667$$

Luego, la variabilidad total incorporada en el experimento es:

$$\hat{\sigma^2} + \hat{\sigma}^2_\alpha + \hat{\sigma}^2_\beta = 0.1299 + 0.07457889 + 0.06593667 = 0.2704156$$

Y puede realizarse una tabla como sigue:

| Fuente | Variabilidad | Porcentaje variabilidad |
|---|---|---|
| Factor máquina | 0.0745 | 27.6% |
| Factor estación | 0.0659 | 24.3% |
| Residuos | 0.1299 | 48% |

Es posible observar que utilizando los factores máquinas y estación se logra explicar el 50% de variabilidad presente en el experimiento.

d) Los estándares de calidad le exijen a la empresa que el promedio de error de ensamblaje no supere los 5 centímetros. Determine si la empresa cumple o no cumple el estándar de calidad.

Respuesta

Estamos interesados en testear:

$$H_0 : \mu \leq 5 \quad H_1 : \mu > 5$$

Es decir, en realizar un test t de una muestra:

```
t.test(error, mu=5, alternative=c("greater"))

	One Sample t-test

data:  error
t = -19.744, df = 26, p-value = 1
alternative hypothesis: true mean is greater than 5
95 percent confidence interval:
 3.032436      Inf
sample estimates:
mean of x
 3.188889
```

Se obtiene un valor p asociado 1, es decir, no se rechaza la hipótesis nula de que la media del error de ensamblaje cometido no supera los 5 centímetros, por lo cual, con los datos obtenidos se podría concluir que la empresa sí estaría cumpliendo el estándar exigido, pero recordar que esto cambiará en cada experimento pues se seleccionaron tanto máquinas como estaciones aleatorias.

**Ejercicio 2: Anova caso mixto**

En una compañía productora de tabacos se desea evaluar la cantidad de los cigarrillos producidos por 2 máquinas diferentes. Con este objetivo, se plantea testear si ambas máquinas difieren en el número de cigarrillos de buena calidad producidos por minuto. Se sospecha que esta cantidad también se ve afectada por el operario que se encuentre de turno, para ello se escogen tres operarios al azar. Cada operario trabaja durante 5 lapsos de un minuto con cada una de las máquinas. Los datos se encuentran en la base de datos `tabaco`.

a) ¿Qué modelo le sugeriría estudiar a la compañía productora de tabacos? Explicite los supuestos y componentes del modelo propuesto.

   Respuesta

   Del ejercicio se desprende que hay dos factores que se cree influirían en la cantidad de cigarrillos de buena calidad producidos, el factor máquina, el cual es fijo pues se tienen en la empresa dos máquinas diferentes, y el factor operario, el cuál es aleatorio pues se elijen 3 operarios al azar.

Así, el modelo propuesto es el siguiente:

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk}$$

Con las siguientes características:

$$\sum_{i}^{2} \alpha_i = 0 \qquad \text{Efecto asociado a la máquina por sobre la cantidad media de cigarrillos}$$

$$\beta_j \overset{\text{i.i.d}}{\sim} N(0, \sigma_\beta^2) \qquad \text{Efecto aleatorio de cada operario}$$

$$\sum_{i}^{2} (\alpha\beta)_{ij} = 0 \qquad \text{Efecto de la interacción de máquina-operario}$$

$$(\alpha\beta)_{ij} \overset{\text{i.i.d}}{\sim} N(0, \frac{a-1}{a}\sigma_{\alpha\beta}^2)$$

$$\epsilon_{ijk} \overset{\text{i.i.d}}{\sim} N(0, \sigma^2)$$

b) Estime las variabilidades de cada componente del modelo planteado. ¿Qué podrían sugerir las cifras obtenidas?

Respuesta

Sabemos que:

$$\sigma_\alpha^2 = \frac{E(MCA) - E(MCAB)}{nb}$$
$$\sigma_\beta^2 = \frac{E(MCB) - E(MCAB)}{na}$$
$$\sigma_{\alpha\beta}^2 = \frac{E(MCAB) - E(MCE)}{n}$$

```
library(tidyverse)

glimpse(tabaco)
Observations: 30
Variables: 3
$ respuesta <dbl> 700, 850, 820, 640, 920, 480, 460, 500, 570, 580, 500, 550...
$ operario  <dbl> 1, 1, 1, 1, 1, 2, 2, 2, 2, 2, 3, 3, 3, 3, 3, 1, 1, 1, 1, 1...
$ maquina   <dbl> 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 1, 2, 2, 2, 2, 2...

cigarros<-tabaco$respuesta
maquina<-factor(tabaco$maquina)
operario<-factor(tabaco$operario)

contrasts(maquina)<-contr.sum

inter<-aov(cigarros~maquina*operario)

anova(inter)[,1:3]
                 Df Sum Sq Mean Sq
maquina           1  25404   25404
operario          2 503215  251608
maquina:operario  2   2817    1408
Residuals        24 106659    4444
```

Y es posible estimar de manera puntual utilizando:

$$\hat{\sigma}_{\alpha}^2 = \frac{MCA - MCAB}{nb} = \frac{25404 - 1408}{5 \cdot 3} = 1599.733$$

$$\hat{\sigma}_{\beta}^2 = \frac{MCB - MCAB}{na} = \frac{251608 - 1408}{5 \cdot 2} = 25020$$

$$\hat{\sigma}_{\alpha\beta}^2 = \frac{MCAB - MCE}{n} = \frac{1408 - 4444}{5} = -607.2$$

Por un lado, es posible interpretar que $\hat{\sigma}_{\alpha}^2 = 1599.733$ se asocia a la variabilidad asociada en la cantidad media de cigarrillos por máquina. Por otro lado, $\hat{\sigma}_{\beta}^2 = 25020$ se asocia a la variabilidad obtenida en el procedimiento de selección aleatoria de operadores y las diferencias en la cantidad media de cigarrillos de calidad producida por operador aleatorio. Note que la estimación de $\sigma_{\alpha\beta}^2$ da negativa, por lo que, pudiera pensarse que este modelo no es el adecuado, sugiriendo que la interacción no logra aportar mayor variabilidad al modelo. Debe realizarse análisis de significancia.

c) Realice el análisis pertinente para complementar sus hallazgos en el partado anterior. Comente.

Respuesta

Siguiendo las intuiciones del apartado anterior, ee desea analizar si existe efecto en la combinación de factores máquina y operario, es decir, si la variabilidad asociada a la componente de la interacción es no nula. El test es el siguiente:

$$H_0 : \sigma_{\alpha\beta}^2 = 0 \qquad H_1 : \sigma_{\alpha\beta}^2 > 0$$

La regla de decisión es:

$$\text{Si} \quad F_{AB}* = \frac{MCAB}{MCE} > F^{1-\alpha}_{(a-1)\cdot(b-1),a\cdot b\cdot(n-1))} \quad \text{Se rechaza la hipótesis nula}$$

```
(a<-length(levels(maquina)))
[1] 2

(b<-length(levels(operario)))
[1] 3

(n<-unique(table(operario, maquina)))  #Caso balanceado
[1] 5

(FAB=anova(inter)[3,3]/anova(inter)[4,3])
[1] 0.3168897

FAB>qf((1-alpha), (a-1)*(b-1), a*b*(n-1))
[1] FALSE
```

Se puede concluir utilizando una confianza del 95%, que la interacción entre los factores máquina y operario no resulta significativa. Se propone un modelo aditivo.

d) En base a lo obtenido en c), ¿existe evidencia para afirmar que el factor máquina y el factor operador inciden significativamente sobre la media de la cantidad de cigarrillos de calidad producidos?

Respuesta

Se debe plantear el modelo aditivo y realizar análisis de significancia de cada factor:

```
aditivo<-aov(cigarros~maquina+operario)

anova(aditivo)[,1:3]
          Df Sum Sq Mean Sq
maquina    1  25404   25404
operario   2 503215  251608
Residuals 26 109476    4211
```

Se realizan los tests de significancia de interés en el modelo aditivo planteado:

$$H_0 : \alpha_i = 0 \ \ \forall \ \ i = 1, 2 \qquad\qquad H_1 : \exists \ \ \alpha_i \neq 0$$

La regla de decisión es:

$$\text{Si} \quad F_A* = \frac{MCA}{MCE} > F^{1-\alpha}_{(a-1),a\cdot b\cdot(n-1)+(a-1)\cdot(b-1)} \quad \text{Se rechaza la hipótesis nula}$$

```
(FA=anova(aditivo)[1,3]/anova(aditivo)[3,3])
[1] 6.033405

FA>qf((1-alpha), (a-1), a*b*(n-1)+(a-1)*(b-1))
[1] TRUE
```

Utilizando un 95% de confianza se rechaza que los efectos asociados a cada nivel del factor máquina sean todos simultáneamente nulos, por lo que, se piensa que el factor máquina sí logra explicar diferencias en términos de la cantidad media de cigarrillos de calidad producidos en un minuto.
Para el factor operario (aleatorio) se realiza el siguiente test de hipótesis:

$$H_0 : \sigma^2_\beta = 0 \qquad\qquad H_1 : \sigma^2_\beta > 0$$

La regla de decisión es:

$$\text{Si} \quad F_B* = \frac{MCB}{MCE} > F^{1-\alpha}_{(b-1)\cdot(b-1),a\cdot b\cdot(n-1)+(a-1)\cdot(b-1)} \quad \text{Se rechaza la hipótesis nula}$$

```
(FB=anova(aditivo)[2,3]/anova(aditivo)[3,3])
[1] 59.75566

FB>qf((1-alpha), (b-1), a*b*(n-1)+(a-1)*(b-1))
[1] TRUE
```

Utilizando un 95% de confianza se rechaza que la variabilidad entre operadores sea nula, es decir, la elección aleatoria de operarios sí genera diferencias en términos de la cantidad media de cigarrillos de calidad producidos en un minuto.

e) Obtenga un intervalo de confianza aproximado para la variabilidad asociada al factor operario. Comente.

Respuesta

Para construir un intervalo de confianza en un modelo aditivo, utilizamos el siguiente procedimiento:

$$gl\frac{\hat{\sigma}^2_\beta}{X^2_{gl,1-\frac{\alpha}{2}}} \leq \sigma^2_\beta \leq gl\frac{\hat{\sigma}^2_\beta}{X^2_{gl,\frac{\alpha}{2}}}$$

Donde $gl = \frac{(na)^2(\hat{\sigma}^2_\beta)^2}{\frac{MCB^2}{b-1}+\frac{MCE^2}{a\cdot b\cdot(n-1)+(a-1)\cdot(b-1)}}$

```
MCB<-anova(aditivo)[2,3]
MCE<-anova(aditivo)[3,3]

(sigmahat<-(MCB-MCE)/(n*a))
[1] 24739.7

glE<-a*b*(n-1)+(a-1)*(b-1)
(gl<-((sigmahat)^{2}*(n*a)^{2})/(MCB^{2}/(b-1) + MCE^2/glE))
[1] 1.933579

(gl<-round(gl)) #Debe aproximarse pues se utiliza como grado de libertad
[1] 2

#El intervalo es:
c(gl*sigma/qchisq(0.975,gl), gl*sigma/qchisq(0.025,gl))
[1]   6706.563 977166.056
```

El intervalo es bastante amplio, esto se debe principalmente a que si bien se sabe que el estimador de la varianza es insesgado, al tomar un valor puntual, la calidad de éste dependará de la cantidad de observaciones y categorías de los factores involucrados. Notar que sólo tenemos $n = 5$.

**Ejercicio 3: Ancova**

La idea del ancova es tratar de reducir la variabilidad inexplicada por los datos, incluyendo predictores categóricos y continuos a la vez.

**Modelo de análisis de la covarianza de un factor (efecto fijo)**

Un modelo que podría plantearse es el siguiente:

$$Y_{ij} = \mu + \tau_i + \gamma X_{ij} + \epsilon_{ij}$$

$$\epsilon_{ij} \overset{\text{i.i.d}}{\sim} N(0, \sigma^2)$$

$$\sum_{i=1}^{r} \tau_i = 0$$

**El problema de este modelo:**

$$\bar{Y}_{..} = \mu + \bar{\gamma}\bar{X}_{..} + \bar{\epsilon}_{..} \Longleftrightarrow \mu = \bar{Y}_{..} - \bar{\gamma}\bar{X}_{..} + \bar{\epsilon}_{..}$$

$\mu$ depende de $\bar{X}$. Una manera de arreglar esto es incorporar la información de $X$ centrado en 0 (es decir, restándole su media):

$$Y_{ij} = \mu + \tau_i + \gamma(X_{ij} - \bar{X}_{..}) + \epsilon_{ij}$$

De este modo, $\bar{Y}_{..} = \mu + \bar{\epsilon}_{..}$
Note que $\mu + \tau_i$ corresponde al intercepto para el nivel $i$ del factor y $\gamma$ corresponde a la pendiente. Es decir, se obtienen $r$ rectas paralelas, con $r$ la cantidad de niveles del factor.

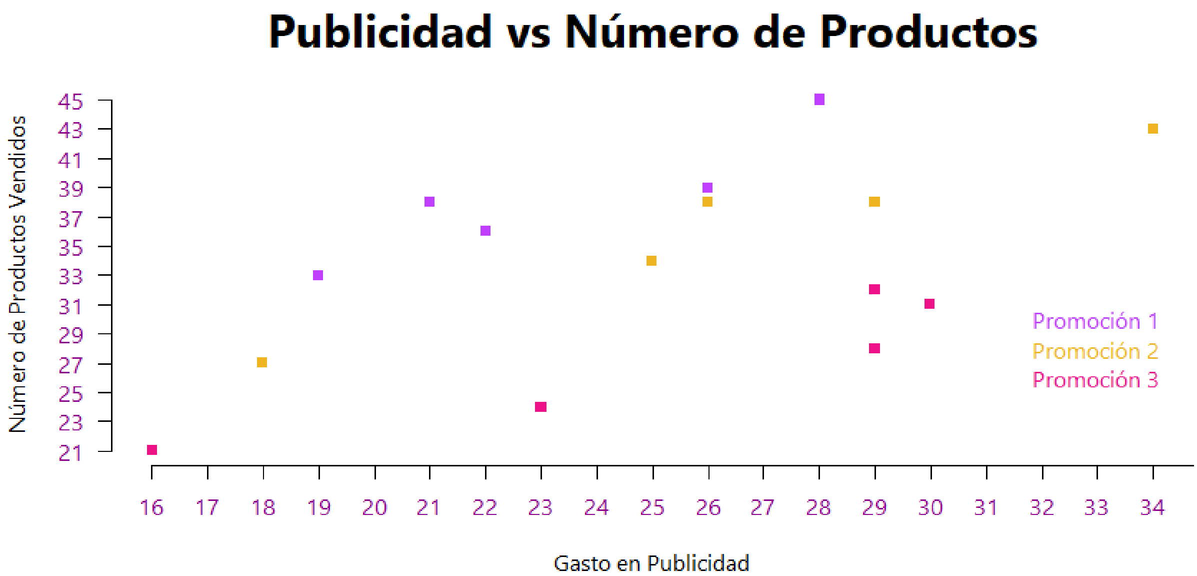
Una compañía estudió los efectos de tres tipos distintos de promociones en sus ventas.

**Promoción 1**: Muestra de productos por los cendedores y con un estante.
**Promoción 2**: Estante adicional en una ubicación regular.
**Promoción 3**: Estantes especiales a un costado, además de los estantes regulares.

15 tiendas fueron seleccionadas por el estudio, y a cada tienda se le asignó uno de los tipos de promociones, con 5 tiendas asignadas a cada una de ellas. El número de productos vendidos durante el tiempo de promoción y el tipo de promoción utilizado en cada caso se encuentra en *promocion.txt*. Se piensa que otra variable para explicar el número de productos vendidos es el monto gastado en publicidad por cada tienda.

a) Evalúe diferencias por promoción y su efecto en el número de productos vendidos.

```
attach(promociones)

#install.packages("extrafont")
library(extrafont)

#font_import()
#loadfonts(device="win")      #Importa tipos de letra de windows
fonts()

par(family ="Nirmala UI")

plot(montopublicidad,numero,pch="",main="Publicidad vs Numero de Productos",
     xlab="Gasto en Publicidad",ylab="Numero de Productos Vendidos", cex.main=2, axes=FALSE)
points(montopublicidad[promocion==1],numero[promocion==1],col="darkorchid1",pch=15)
points(montopublicidad[promocion==2],numero[promocion==2],col="goldenrod2",pch=15)
points(montopublicidad[promocion==3],numero[promocion==3],col="deeppink2",pch=15)
text(x=33, y= 30, labels="Promocion 1", col="darkorchid1", cex=1)
text(x=33, y= 28, labels="Promocion 2", col="goldenrod2", cex=1)
text(x=33, y= 26, labels="Promocion 3", col="deeppink2", cex=1)

axis(1, at=seq(16,35, by=1), col.axis="darkmagenta", las=1)
axis(2, at=seq(21,45, by=2), col.axis="darkmagenta", las=1)
```



b) Plantee el modelo, obtenga las pendientes asociadas a cada promoción e interprete.

Respuesta

El modelo corresponde a un caso de análisis de covarianzas de un factor de efectos fijos:

$$Y_{ij} = \mu + \tau_i + \gamma(X_{ij} - \bar{X}_{..}) + \epsilon_{ij}$$

Con

$$\epsilon_{ij} \overset{\text{i.i.d}}{\sim} N(0, \sigma^2)$$

$$\sum_{i=1}^{r} \tau_i = 0$$

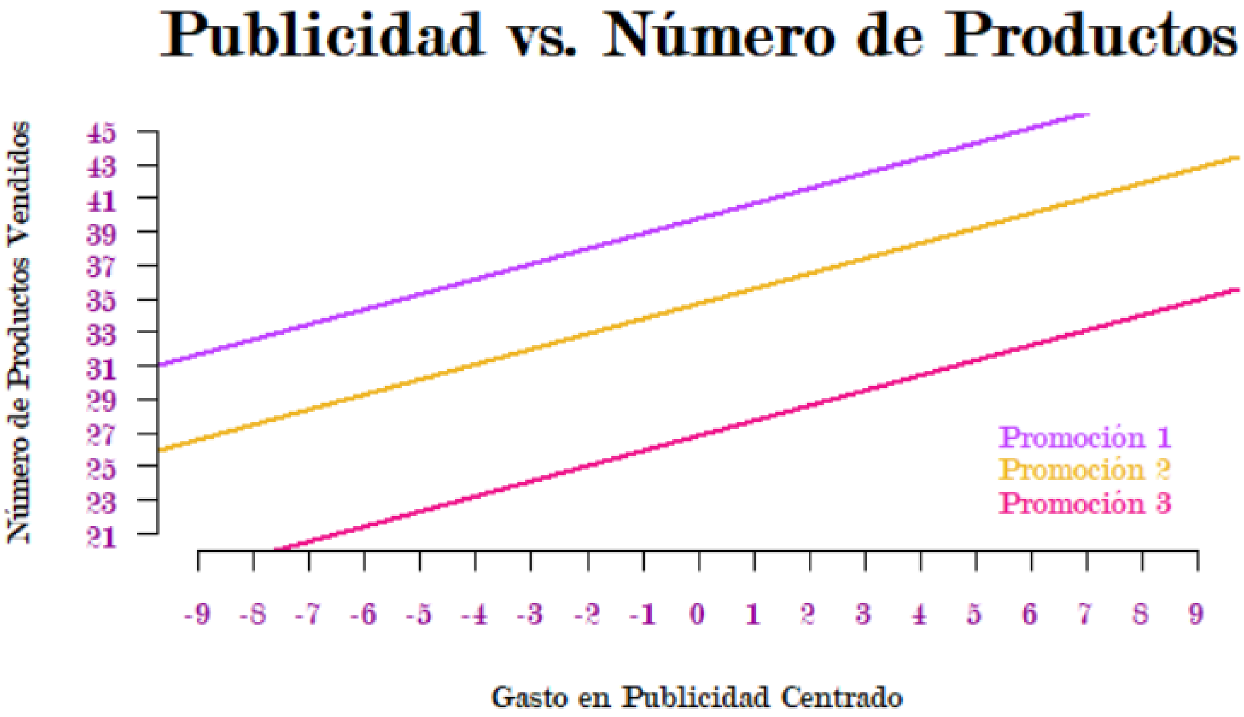
$\tau_i$ corresponde al efecto asociado a cada promoción por sobre el número medio de productos vendidos. $X_{ij}$ corresponde al $j$ ésima monto invertido en la promoción $i$. $\tau_i + \mu$ es la pendiente para la promoción $i$.

```
promocion<-as.factor(promocion)
publicidad.c<-montopublicidad-mean(montopublicidad)
contrasts(promocion)<-contr.sum

modelo<-lm(numero~publicidad.c+promocion)

modelo$contrasts
$promocion
  [,1] [,2]
1    1    0
2    0    1
3   -1   -1

modelo$coefficients
(Intercept) publicidad.c   promocion1   promocion2
 33.8000000    0.8985594    6.0174070    0.9420168

(tauprom3<--(6.0174070+0.9420168))
[1] -6.959424
```

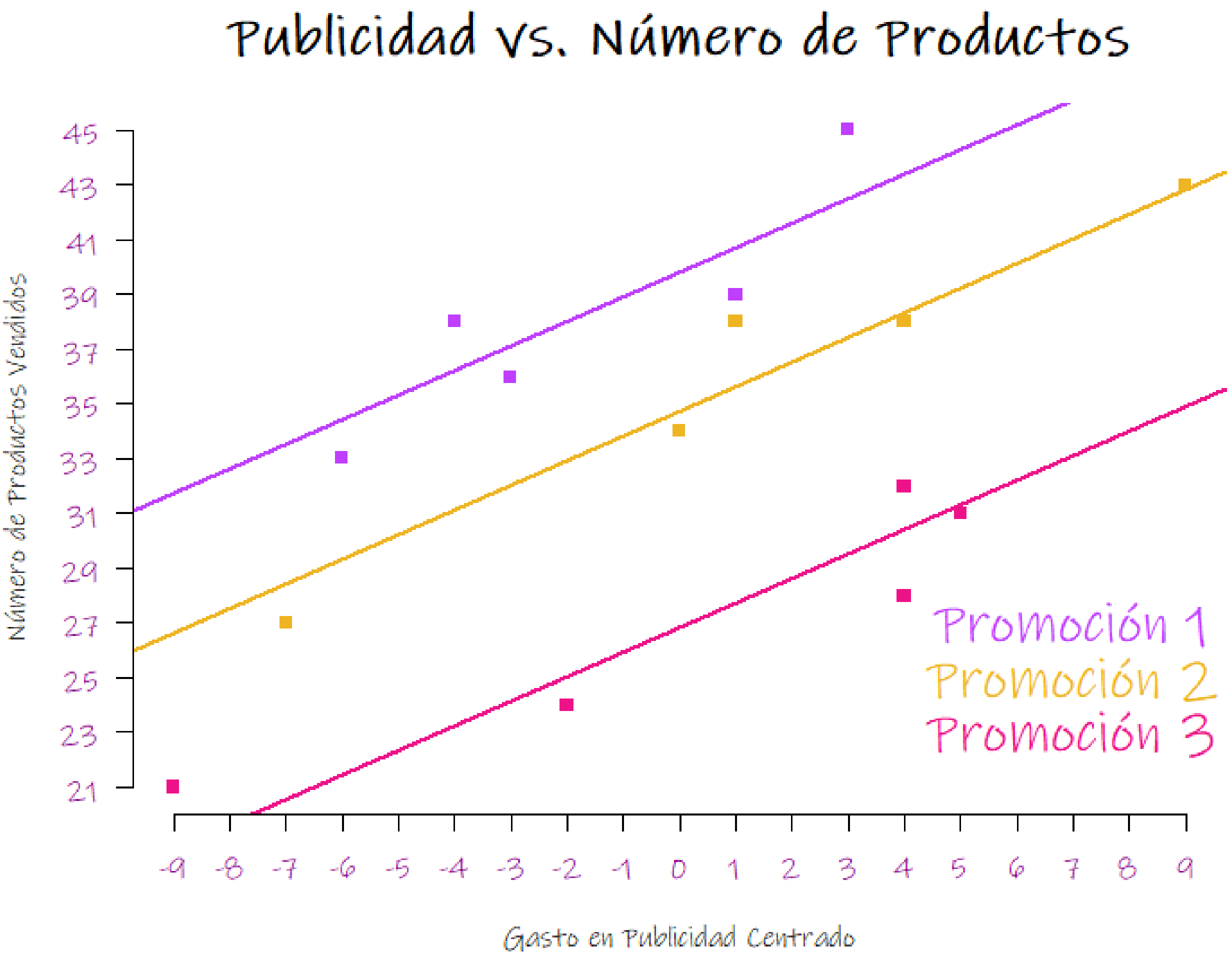
Las pendientes para cada promoción son:

```
#Intercepto para la promocion 1
(intercepto.1<-coef(modelo)[1]+coef(modelo)[3])
(Intercept)
   39.81741

#Intercepto para la promocion 2
(intercepto.2<-coef(modelo)[1]+coef(modelo)[4])
(Intercept)
   34.74202

#Intercepto para la promocion 3
(intercepto.3<-coef(modelo)[1]+tauprom3)
(Intercept)
   26.84058

#Grafico
par(family = "Modern No. 20")
plot(publicidad.c,numero,pch="",main="Publicidad vs. Numero de Productos",
     xlab="Gasto en Publicidad Centrado",ylab="N mero de Productos Vendidos", axes=FALSE,
```

```
        cex.main=2)
    abline(intercepto.1,modelo$coefficients[2],col="darkorchid1", lwd=2)
    abline(intercepto.2,modelo$coefficients[2],col="goldenrod2", lwd=2)
    abline(intercepto.3,modelo$coefficients[2],col="deeppink2", lwd=2)
    text(x=7, y= 27,labels="Promocion 1", col="darkorchid1", cex=1)
    text(x=7, y= 25,labels="Promocion 2", col="goldenrod2", cex=1)
    text(x=7, y= 23,labels="Promocion 3", col="deeppink2", cex=1)
    axis(1, at=seq(-9,9, by=1), col.axis="darkmagenta", las=1)
    axis(2, at=seq(21,45, by=2), col.axis="darkmagenta", las=1)
```




```
par(family = "Ink Free")
plot(publicidad.c,numero,pch="",main="Publicidad vs. N mero de Productos",
     xlab="Gasto en Publicidad Centrado",ylab="N mero de Productos Vendidos", axes=FALSE, cex.main=2)
points(publicidad.c[promocion==1],numero[promocion==1],col="darkorchid1",pch=15)
points(publicidad.c[promocion==2],numero[promocion==2],col="goldenrod2",pch=15)
points(publicidad.c[promocion==3],numero[promocion==3],col="deeppink2",pch=15)
abline(intercepto.1,modelo$coefficients[2],col="darkorchid1", lwd=2)
abline(intercepto.2,modelo$coefficients[2],col="goldenrod2", lwd=2)
abline(intercepto.3,modelo$coefficients[2],col="deeppink2", lwd=2)
text(x=7, y= 27,labels="Promoci n 1", col="darkorchid1", cex=2)
text(x=7, y= 25,labels="Promoci n 2", col="goldenrod2", cex=2)
text(x=7, y= 23,labels="Promoci n 3", col="deeppink2", cex=2)
axis(1, at=seq(-9,9, by=1), col.axis="darkmagenta", las=1)
axis(2, at=seq(21,45, by=2), col.axis="darkmagenta", las=1)
```


Publicidad vs. Número de Productos
Número de Productos Vendidos
45
43
41
39
37
35
33
31
29
27
25
23
21
-9 -8 -7 -6 -5 -4 -3 -2 -1 0 1 2 3 4 5 6 7 8 9
Gasto en Publicidad Centrado
Promoción 1
Promoción 2
Promoción 3