maximiliano.vaca@uabc.edu.mx Profesor: Andrés García Medina andres.garcia.medina@uabc.edu. Fecha de entrega: jueves 14 de m	mx
Resuelva los siguientes problem :D Ejercicio 2	as conceptuales:
Considere los datos en R denom	inados InsectSprays, los cuales contienen recuentos de insectos en cada una de varias parcelas. Cada una de las parcelas ha sido rociada con un insecticida distinto. Un modelo para estos datos podría ser: $y_i=\mu+eta_j$ $\mu=NOmedia$ $\mu\simeta_0$
falta el epsilon mejor:	$y_i = \mu + eta_j + \epsilon_i$
si la parcela i ha sido rociada con data (InsectSprays) head (InsectSprays)	$y_i = lpha + eta_j + \epsilon_i$ n el spray j
<pre>#descripcion de los datos str(InsectSprays) A data.frame: 6 × 2 count spray <dbl> <fct></fct></dbl></pre>	
 1 10 A 2 7 A 3 20 A 4 14 A 	
<pre>5 14 A 6 12 A 'data.frame': 72 obs. of \$ count: num 10 7 20 14 1 \$ spray: Factor w/ 6 level</pre>	
as.factor(InsectSprays\$spr	ene el conjunto de datos? Se puede apoyar en el comando $as.\ factor$ ay) A · B · B · B · B · B · B · B · B · B ·
▶ Levels:• la variable spray es 1 facto• pairs(InsectSprays)	r en 6 niveles
count occupation occu	
 ω - 00000 0 + - 00000 0 0 - 000000 	spray
(b) Ajuste un modelo lineal prop	ouesto para determinar si el tipo de insecticida determina el conteo de insectos, y en dado caso en que medida. ¿Cual insecticida tiene mayor impacto en la disminución del numero de mosquitos?
si la parcela i ha sido rociada cor i : #ajuste del modelo	$y_i = lpha + eta_j + \epsilon_i$ n el spray j
<pre>model_1 <- lm(count ~ spra summary(model_1) coef(model_1) Call: lm(formula = count ~ spray,</pre>	
sprayB 0.8333 1. sprayC -12.4167 1. sprayD -9.5833 1. sprayE -11.0000 1. sprayF 2.1667 1.	6011 0.520 0.604 6011 -7.755 7.27e-11 *** 6011 -5.985 9.82e-08 *** 6011 -6.870 2.75e-09 *** 6011 1.353 0.181
Residual standard error: 3. Multiple R-squared: 0.7244 F-statistic: 34.7 on 5 and	01 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 922 on 66 degrees of freedom , Adjusted R-squared: 0.7036 66 DF, p-value: < 2.2e-16 33333332 sprayC: -12.4166666666667 sprayD: -9.5833333333 sprayE: -11 sprayF: 2.1666666666666666666666666666
(Intercept): 14.5 sprayC: -12.4166	de mayor a menor pef(model_1)),decreasing = T)] 666666667 sprayE:-11 sprayD:-9.58333333333333333333333333333333333333
	$count = 14.5 + 0.83 * sprayB - 12.42 * sprayC - 9.58 * sprayD - 11.0 * sprayE + 2.17 * sprayF$ do al valor pequeño del $p-value$ ne el mayor impacto en la disminución del numero de mosquitos, ademas de con un nivel de significancia de $***$, aunque no muy por detrás queda el $incecticida\ C$
<pre>X <- model.matrix(count ~ y <- InsectSprays\$count dim(X) head(X) #tail(X)</pre>	spray, data=InsectSprays)
72 · 6 A matrix: 6 × 6 of to the contract of	sprayD sprayE sprayF 0 0 0
2 1 0 0 3 1 0 0 4 1 0 0 5 1 0 0	
(c) Determine el tipo impacto que Hint: Use la función relevel #cambiamos el nivel de ref	0 0 0 ue tiene el spray A sobre el conteo de mosquitos usando como base el nivel F. Ferencia a F
	evel(InsectSprays\$spray, ref = "F")
<pre>lm(formula = count ~ spray, Residuals: Min 10 Median 30 -8.333 -1.958 -0.500 1.667 Coefficients:</pre>	Max
Estimate Std. E (Intercept) 16.667 1 sprayA -2.167 1 sprayB -1.333 1 sprayC -14.583 1 sprayD -11.750 1	rror t value Pr(> t) .132 14.721 < 2e-16 *** .601 -1.353 0.181 .601 -0.833 0.408 .601 -9.108 2.79e-13 *** .601 -7.339 4.04e-10 ***
Signif. codes: 0 '***' 0.0 Residual standard error: 3. Multiple R-squared: 0.7244	.601 -8.223 1.05e-11 *** 01 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 922 on 66 degrees of freedom , Adjusted R-squared: 0.7036 66 DF, p-value: < 2.2e-16
 otra observacion es que, en 	encia a F , el $sprayA$ tiene un impacto de -2.167 , aunque con un $p-value$ de 0.181 , no creo que sea una estimacion significativa este caso, el $incecticida\ C$ tiene el mayor impacto en la disminución del numero de mosquitos, aunque no por mucho por delante del $incecticida\ E$, el cual habiamos obtenido anteriormente. a través de los cuantiles de los residuales con apoyo de la función qnorm. Compare su gráfico con los que se obtienen por defecto mediante la función lm.
<pre>"Este gráfico compara los cuantil "Este gráfico compara los cuantil "guardamos residuales residuals <- resid(model_1 #grafico Q-Q de los residuales</pre>	
<pre>qqnorm(residuals) qqline(residuals) #linea o</pre>	
ilies –	
Sample Quan	
-2 -1	T T T T T T T T T T T T T T T T T T T
plot(model) muestra: 1. Residuales contra ajustes 2. Gráfico Q-Q de los residuale	
<pre>3. Escala-ubicación (que es la r 4. Residuales contra orden de #plot(model) plot(model_1, which = 2)</pre>	raíz cuadrada de los residuales estandarizados contra ajustes) influencia
Norm	al Q-Q 069 700 08 0°
tandardized residuals 0 1	
2 0 0 0 0 0 0 0 0 0 0	
	1 1 2 al Quantiles t ~ spray)
(e) Omita los posibles datos atip # Calcular los residuales standardized_residuals < #analizamos	
<pre>#omitimos residuales atipa outliers <- abs(standardiz outliers</pre>	icos red_residuals) > 1 # mayor a 2 no sirvió mucho, tomando como referencia los graficos
-2.21930160690107 · 2.485617799 1: TRUE 2: TRUE 3: TRUE 4: FALSE	- InsectSprays[!outliers,]
]: #modelo sin datos atipicos	ount ~ spray, data = InsectSprays_no_outliers)
	data = InsectSprays_no_outliers) 3Q Max 8571 3.0000
Estimate Std. E (Intercept) 14.4000 0. sprayA -0.4000 0. sprayB 1.7429 0. sprayC -12.7636 0. sprayD -10.1273 0.	rror t value Pr(> t) 6909 20.843 < 2e-16 *** 9355 -0.428 0.6709 9046 1.927 0.0602 . 8332 -15.318 < 2e-16 *** 8332 -12.154 5.79e-16 ***
Signif. codes: 0 '***' 0.0 Residual standard error: 1. Multiple R-squared: 0.9392	8223 -13.255 < 2e-16 *** 01 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 545 on 46 degrees of freedom , Adjusted R-squared: 0.9326 46 DF, p-value: < 2.2e-16
	y despues de quitar outliers (esto se estuvo haciendo repetidamente para determinar el valor maximo permitido de std residuals)
Norm	al Q-Q 210 90
dardized residuals 0 1	
Standard	
-2 -1 Theoretical Im(coun	0 1 2 al Quantiles t ~ spray)
Norm	al Q-Q 069 700 08 00 00 00 00 00 00 00 00 00 00 00 0
Standardized residuals 0 1	
lm(coun	nejor en terminos de R2 ajustada y el estadistico F?
<pre>Call: lm(formula = count ~ spray, Residuals: Min 10 Median</pre>	<pre>data = InsectSprays_no_outliers) 30 Max</pre>
(Intercept) 14.4000 0. sprayA -0.4000 0. sprayB 1.7429 0.	rror t value Pr(> t) 6909 20.843 < 2e-16 *** 9355 -0.428 0.6709 9046 1.927 0.0602 .
<pre>sprayC -12.7636 0. sprayD -10.1273 0. sprayE -10.9000 0 Signif. codes: 0 '***' 0.0 Residual standard error: 1.</pre>	8332 -15.318 < 2e-16 *** 8332 -12.154 5.79e-16 *** 8233 -13.255 < 2e-16 *** 01 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 545 on 46 degrees of freedom
Multiple R-squared: 0.9392	, Adjusted R-squared: 0.9326 46 DF, p-value: < $2.2e-16$ usted R-squared: 0.7036
F-statistic: 34.7 on 5 and 66 DF, No Outliers: Multiple R-squared: 0.9392, Adj F-statistic: 142.1 on 5 and 46 DF	usted R-squared: 0.9326
Describimos el mejor modelo d	
	to es: $count = 14.4 - 0.4 * sprayA + 1.7429 * sprayB - 12.7636 * sprayC - 10.1273 * sprayD - 10.9000 * sprayE$ do al valor pequeño del $p-value$ ne el mayor impacto en la disminución del numero de mosquitos, ademas de con un nivel de significancia de $***$, aunque no muy por detrás queda el $incecticida$ E
	ltados desarrolle sus conclusiones y sugerencias para un hipotético agricultor nero por utilizar el insecticida C, y como segunda opcion tener el E, ya que en todos los modelos estos demuestran ser los que tienen mas infuencia, ninguno muy por delante del otro, por lo que la segunda opcion tambien
	rar ambos modelos a través del criterio de Akaike . Utilice la función AIC y una implementación propia. ¿Es justa su comparación?
 Mi sugerencia sería irse prin viable. Ejercicio 3 	AIC=-2l+2p
 Mi sugerencia sería irse prin viable. Ejercicio 3 En relación al ejercicio 2, compa aic_1 <- AIC(model_1) aic_2 <- AIC(model_no_outle) 	
• Mi sugerencia sería irse prin viable. Ejercicio 3 En relación al ejercicio 2, compa 3: aic_1 <- AIC(model_1) aic_2 <- AIC(model_no_outle print(paste("AIC model_no_outle paste("AIC model_no_ou	", aic_1)) v: ", aic_2)) 70088124" 7900477599"
• Mi sugerencia sería irse prin viable. Ejercicio 3 En relación al ejercicio 2, compa aic_1 <- AIC(model_1) aic_2 <- AIC(model_no_outle) print(paste("AIC model_ne_outle) print(paste("AIC model_ne_outle) [1] "AIC model_1: 408.8493 [1] "AIC model_n_o: 200.42	", aic_1)) :: ", aic_2)) 70088124" 7900477599" {