PS 22-23 S2 (28-6-2023)

1. Disponemos de los datos muestrales de la EES (Encuesta de Estructura Salarial) generados por el Instituto Nacional de Estadística (INE) en 2018; en particular, hemos seleccionado dos de las variables de estudio: Salario Base (SALBASE) medido en euros, y años de antigüedad (ANOANTI); por otra parte, disponemos de los dos siguientes análisis realizados con RCommander para dos regiones diferentes:

MODELO 1:

```
lm(formula = SALBASE ~ ANOANTI, data = región1)

Residuals:
   Min     1Q Median     3Q    Max
 -1895   -497   -190    219 164847

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 1237.1560     4.1989  294.64   <2e-16 ***
ANOANTI       15.4639     0.2848   54.29   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1351 on 216724 degrees of freedom
Multiple R-squared:  0.01342,	Adjusted R-squared:  0.01341
F-statistic:  2947 on 1 and 216724 DF,  p-value: < 2.2e-16
```

MODELO 2:

```
lm(formula = SALBASE ~ ANOANTI, data = región2)

Residuals:
   Min     1Q Median     3Q    Max
 -2343   -652   -322    187 116242

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 1351.9714    12.6672  106.73   <2e-16 ***
ANOANTI       23.2237     0.8975   25.88   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1625 on 34267 degrees of freedom
Multiple R-squared:  0.01917,	Adjusted R-squared:  0.01914
F-statistic: 669.6 on 1 and 34267 DF,  p-value: < 2.2e-16
```

(2,5 puntos) Sabiendo que, en nuestra región de interés, la acumulación de años de antigüedad se premia con un mayor incremento del salario base, seleccionar el modelo de nuestra región, indicando la recta de regresión lineal simple obtenida, interpretar ambos coeficientes (intercepto y pendiente) y obtener el coeficiente de correlación. **(máximo 150 palabras)**

MODELO 2:

```
lm(formula = SALBASE ~ ANOANTI, data = región2)

Residuals:
   Min     1Q Median     3Q    Max
 -2343   -652   -322    187 116242

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 1351.9714    12.6672  106.73   <2e-16 ***
ANOANTI       +23.2237     0.8975   25.88   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1625 on 34267 degrees of freedom
Multiple R-squared:  0.01917,    Adjusted R-squared:  0.01914
F-statistic: 669.6 on 1 and 34267 DF,  p-value: < 2.2e-16
```

- $\widehat{SALBASE}_i = 1351'97 + 23'2237\ ANOANTI_i$ → Recta regresión
- $b_0 = 1351'97$ $\quad \hat{y} = \widehat{SALBASE} = 1351'97$ → Salario esperado para un individuo que tiene 0 años de antigüedad
  $(X=0)$
  ANOANTI $=0$
- $b_1 = 23'2237$ → Si aumentamos 1 año la antigüedad se espera que el salario aumente en $23'2237$ € en promedio
  $b_1 > 0$
- $r_{xy} = +\sqrt{R^2} = +\sqrt{0'01917} = +0'1384$

2. En un estudio demoscópico realizado en paralelo a la EES de 2018, se concluyó que un 20% de los trabajadores españoles había tenido algún tipo de cambio relacionado con su jornada laboral, mientras un 10% de los trabajadores había tenido cambios en su tipo de contrato; por otra parte, un 7.5% de los trabajadores tuvieron cambios de ambos tipos, es decir, de jornada laboral y de contrato. Si seleccionamos un trabajador al azar, se pide calcular:

(2,5 puntos) Probabilidad de que el trabajador no sufra ninguno de estos dos cambios en sus condiciones (jornada laboral, tipo de contrato). **(máximo 150 palabras)**

JL → cambios jornada laboral

TC → " tipo contrato

$p(JL) = 0{'}2$

$p(TC) = 0{'}1$

$p(JL \cap TC) = 0{'}075$

$$\boxed{p(\overline{JL} \cap \overline{TC})} = 1 - p(\overline{\overline{JL} \cap \overline{TC}}) =$$

$$= 1 - p(\cancel{\overline{\overline{JL}}} \;\overline{\cap}\; \cancel{\overline{\overline{TC}}}) =$$

$$= 1 - \underbrace{p(JL \cup TC)}_{0{'}225} = 1 - 0{'}225 = \boxed{0{'}775}$$

$$p(JL \cup TC) = p(JL) + p(TC) - p(JL \cap TC) = 0{'}2 + 0{'}1 - 0{'}075 = 0{'}225$$

$$p(A \cup B) = p(A) + p(B) - p(A \cap B)$$

3. A continuación, se muestran algunos estadísticos descriptivos de nuestra variable de estudio, Salario Base (SALBASE), medida en euros:

		Q_1	Me	Q_3	
Mean	sd	25%	50%	75%	n
1588.24	1641.05	921.63	1213.34	1763	34269

Se asume que esta variable sigue una distribución Normal con los parámetros obtenidos de la muestra.

(2,5 puntos) Se considera que el valor de una observación es atípicamente alto cuando este es mayor que: Q_3 + 1.5RIQ, siendo RIQ== Rango intercuartítico (Q_3 y RIQ son los valores muestrales).

Entonces, si seleccionamos un trabajador al azar, ¿cuál es la probabilidad de que su salario base sea atípicamente alto? **(máximo 150 palabras)**

$$X \sim N(\mu = 1588'24,\ \sigma = 1641'05)$$

$$p(X > \underline{\underline{LS}})$$

$$\{LS = Q_3 + 1'5(Q_3 - Q_1) = 1763 + 1'5(1763 - 921'63) = 3025'055\}$$

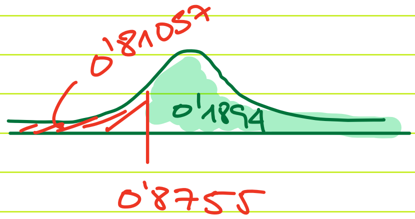
$$\boxed{p(X > 3025'055)} = p\left(Z > \frac{3025'055 - 1588'24}{1641'05}\right) =$$

$$= p(Z > 0'8755) = 1 - p(Z \le 0'88) =$$

$$= 1 - 0'81057 = 0'1894$$

DISTRIBUCIÓN NORMAL ESTÁNDAR

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0	0,50000	0,50399	0,50798	0,51197	0,51595	0,51994	0,52392	0,52790	0,53188	0,53586
0,1	0,53983	0,54380	0,54776	0,55172	0,55567	0,55962	0,56356	0,56749	0,57142	0,57535
0,2	0,57926	0,58317	0,58706	0,59095	0,59483	0,59871	0,60257	0,60642	0,61026	0,61409
0,3	0,61791	0,62172	0,62552	0,62930	0,63307	0,63683	0,64058	0,64431	0,64803	0,65173
0,4	0,65542	0,65910	0,66276	0,66640	0,67003	0,67364	0,67724	0,68082	0,68439	0,68793
0,5	0,69146	0,69497	0,69847	0,70194	0,70540	0,70884	0,71226	0,71566	0,71904	0,72240
0,6	0,72575	0,72907	0,73237	0,73565	0,73891	0,74215	0,74537	0,74857	0,75175	0,75490
0,7	0,75804	0,76115	0,76424	0,76730	0,77035	0,77337	0,77637	0,77935	0,78230	0,78524
0,8	0,78814	0,79103	0,79389	0,79673	0,79955	0,80234	0,80511	0,80785	0,81057	0,81327
0,9	0,81594	0,81859	0,82121	0,82381	0,82639	0,82894	0,83147	0,83398	0,83646	0,83891
1	0,84134	0,84375	0,84614	0,84849	0,85083	0,85314	0,85543	0,85769	0,85993	0,86214
1,1	0,86433	0,86650	0,86864	0,87076	0,87286	0,87493	0,87698	0,87900	0,88100	0,88298
1,2	0,88493	0,88686	0,88877	0,89065	0,89251	0,89435	0,89617	0,89796	0,89973	0,90147
1,3	0,90320	0,90490	0,90658	0,90824	0,90988	0,91149	0,91308	0,91466	0,91621	0,91774
1,4	0,91924	0,92073	0,92220	0,92364	0,92507	0,92647	0,92785	0,92922	0,93056	0,93189
1,5	0,93319	0,93448	0,93574	0,93699	0,93822	0,93943	0,94062	0,94179	0,94295	0,94408
1,6	0,94520	0,94630	0,94738	0,94845	0,94950	0,95053	0,95154	0,95254	0,95352	0,95449
1,7	0,95543	0,95637	0,95728	0,95818	0,95907	0,95994	0,96080	0,96164	0,96246	0,96327
1,8	0,96407	0,96485	0,96562	0,96638	0,96712	0,96784	0,96856	0,96926	0,96995	0,97062
1,9	0,97128	0,97193	0,97257	0,97320	0,97381	0,97441	0,97500	0,97558	0,97615	0,97670
2	0,97725	0,97778	0,97831	0,97882	0,97932	0,97982	0,98030	0,98077	0,98124	0,98169
2,1	0,98214	0,98257	0,98300	0,98341	0,98382	0,98422	0,98461	0,98500	0,98537	0,98574
2,2	0,98610	0,98645	0,98679	0,98713	0,98745	0,98778	0,98809	0,98840	0,98870	0,98899
2,3	0,98928	0,98956	0,98983	0,99010	0,99036	0,99061	0,99086	0,99111	0,99134	0,99158
2,4	0,99180	0,99202	0,99224	0,99245	0,99266	0,99286	0,99305	0,99324	0,99343	0,99361
2,5	0,99379	0,99396	0,99413	0,99430	0,99446	0,99461	0,99477	0,99492	0,99506	0,99520
2,6	0,99534	0,99547	0,99560	0,99573	0,99585	0,99598	0,99609	0,99621	0,99632	0,99643
2,7	0,99653	0,99664	0,99674	0,99683	0,99693	0,99702	0,99711	0,99720	0,99728	0,99736
2,8	0,99744	0,99752	0,99760	0,99767	0,99774	0,99781	0,99788	0,99795	0,99801	0,99807
2,9	0,99813	0,99819	0,99825	0,99831	0,99836	0,99841	0,99846	0,99851	0,99856	0,99861
3	0,99865	0,99869	0,99874	0,99878	0,99882	0,99886	0,99889	0,99893	0,99896	0,99900
3,1	0,99903	0,99906	0,99910	0,99913	0,99916	0,99918	0,99921	0,99924	0,99926	0,99929
3,2	0,99931	0,99934	0,99936	0,99938	0,99940	0,99942	0,99944	0,99946	0,99948	0,99950
3,3	0,99952	0,99953	0,99955	0,99957	0,99958	0,99960	0,99961	0,99962	0,99964	0,99965
3,4	0,99966	0,99968	0,99969	0,99970	0,99971	0,99972	0,99973	0,99974	0,99975	0,99976



4. Supongamos que, en una determinada región, el 15% de los trabajadores tienen un salario base superior a 1100€. Si seleccionamos al azar una muestra de 120 trabajadores ~~hogares~~ en esta región:

(2,5 puntos) Obtenga la media y la desviación típica de la proporción de trabajadores con un salario base superior a 1100€ en dicha muestra. **(máximo 150 palabras)**

Población

$p_{>1100€} = 0'15$

$n = 120$

$$\boxed{E(\hat{p}) = p = 0'15}$$

$$Var(\hat{p}) = \frac{p(1-p)}{n} = 0'00106$$

$$\sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}} =$$

$$= \sqrt{\frac{0'15 \cdot (1-0'15)}{120}} = 0'0325$$

$$\boxed{\sigma_{\hat{p}} = 0'0325}$$