

Práctico 9: Test de Bondad de Ajuste

Test de Ajuste de Kolmogorov-Smirnov-Lilliefors

1. Los siguientes datos corresponden a una muestra aleatoria simple de la velocidad media del viento en la ciudad de Colonia durante el año 2013:

| | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 5.3 | 5.1 | 4.8 | 4.9 | 5.3 | 5.2 | 5.5 | 5.8 | 5.6 | 5.2 |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|

- Graficar la distribución uniforme en el intervalo $[4.7, 5.9]$.
 - Dibujar en el mismo gráfico la distribución empírica de los datos de viento.
 - Determinar en el gráfico la máxima discrepancia D entre las dos funciones dibujadas. Indicar su valor.
 - Realizar en **R** un test de Kolmogorov-Smirnov (KS) para determinar si es posible asumir que la distribución de los datos es uniforme en el intervalo $[4.7, 5.9]$. Usar el comando `ks.test`.
2. Los siguientes datos representan las variaciones de población (en porcentajes) en un país por regiones. Se pretende estudiar si las variaciones son comparables.

| | | | | | | | | | |
|------------|------|-------|-------|-------|-------|------|------|-----|-----|
| rural | 1.1 | -21.7 | -16.3 | -11.3 | -10.4 | -7 | -2 | 1.9 | 6.2 |
| no rurales | -2.4 | 9.9 | 14.2 | 18.4 | 20.1 | 23.1 | 70.4 | | |

- Graficar en los mismos ejes las distribuciones empíricas de los datos correspondientes a las zonas rurales y no rurales.
 - Determinar en el gráfico la máxima discrepancia D entre las dos funciones dibujadas. Indicar su valor.
 - Realizar en **R** un test KS de dos muestras para determinar si es posible asumir que las distribuciones de los datos son iguales. Usar el comando `ks.test`.
3. Los siguientes datos corresponden a una muestra aleatoria simple de la concentración de hierro en una aleación.

| | | | | | | | |
|-------|----------|----------|----------|----------|----------|----------|----------|
| X_1 | X_2 | X_3 | X_4 | X_5 | X_6 | X_7 | X_8 |
| 49.83 | 45.08 | 49.33 | 50.47 | 51.29 | 49.85 | 51.93 | 47.24 |
| X_9 | X_{10} | X_{11} | X_{12} | X_{13} | X_{14} | X_{15} | X_{16} |
| 45.77 | 49.89 | 49.04 | 51.23 | 52.63 | 50.22 | 49.11 | 51.22 |

- Realizar en **R** un test KS para determinar si se puede asumir que los datos son $\mathcal{N}(50, 4)$. Determinar si el p -valor del test es mayor que 0.05.
- Estimar la media y la varianza de los datos y estandarizar los datos.
- Hallar la máxima discrepancia D entre la distribución empírica de los datos estandarizados y la distribución normal estándar.
- Realizar la prueba de normalidad de Lilliefors para los datos normalizados con el comando `LcKS` de **R**.

Observación: Para usar el comando `LcKS` hay que instalar primero el paquete `KScorrect` y cargarlo a la librería (solo se instala una vez). Ejemplo:

```
install.packages('KScorrect')
library('KScorrect')
LcKS(datos, 'pnorm')
```

4. Sea la muestra aleatoria simple de tiempos de vida de ciertos microorganismos.

| | | | | | | | | | |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|----------|
| X_1 | X_2 | X_3 | X_4 | X_5 | X_6 | X_7 | X_8 | X_9 | X_{10} |
| 0.85 | 0.70 | 1.16 | 3.95 | 0.94 | 4.04 | 0.56 | 3.93 | 0.25 | 1.45 |

- (a) Se quiere estudiar si se puede considerar que los datos son exponenciales de parámetro $\lambda = 0.25$. Para eso se determina el estadístico D que se presenta en el dato séptimo más pequeño. Determinar D .
- (b) Aplicar en \mathbf{R} un test KS y comparar el valor del estadístico con el obtenido en la parte anterior.
- (c) Queremos estudiar ahora si los datos son exponenciales. Realizar en \mathbf{R} un test de Lilliefors para ver si los datos pueden considerarse exponenciales.

Test de ajuste Chi-cuadrado

5. Se considera un experimento genético bifactorial con los alelos (A/a) y (B/b) . Un modelo supone que el alelo A se presenta con una probabilidad $p_1 = 0,2$, que el B se presenta con una probabilidad $p_2 = 0,3$ y que ambos sitios segregan independientemente.

- (a) Completar la tabla de los valores esperados de los distintos genotipos si se hacen 1000 cruzas.

| | | | |
|------|---------|---------|---------|
| | AA | Aa | aa |
| BB | $e_1 =$ | $e_2 =$ | $e_3 =$ |
| Bb | $e_4 =$ | $e_5 =$ | $e_6 =$ |
| bb | $e_7 =$ | $e_8 =$ | $e_9 =$ |

- (b) Comparar mediante el estadístico χ^2 la tabla de los valores esperados con la de los observados

| | | | |
|------|------|------|------|
| | AA | Aa | aa |
| BB | 3 | 27 | 61 |
| Bb | 18 | 123 | 281 |
| bb | 18 | 152 | 317 |

para ver si el modelo propuesto es razonable.