Probabilidad y Estadística en Física Experimental

Práctica 5

Intervalos de Confianza. Clasificadores. Test de Hipótesis.

Ejercicios

1. Una nueva partícula del tipo *charm* es vista en un detector de emulsión expuesto a un haz de neutrinos. Midiendo la energía de las partículas secundarias producidas por esta partícula, se determina que para este particular evento, la partícula encantada vivió 3x10^{-13} segundos en su sistema en reposo antes de decaer ¿Cuáles son lo limites, a un nivel de confianza de 90 % (simétrico), para la vida media de esta partícula?

   Rta: 1 10^{-13} seg , 5.85 10^{-12} seg

2. Se obtienen 6 medidas independientes de una variable con valores 10.7, 9.7, 13.3, 10.2, 8.9, 11.6, con incerteza conocida $\sigma = 2$. Asumiendo que siguen una distribución normal, encuentre el intervalo de confianza simétrico para μ correspondiente a (a) 90 % (b) 95 % (c) 99 % .

   Rta: 9.39 - 12.08 , 9.13 - 12.33, 8.63 - 12.84

3. El bosón W es una partícula elemental intermediaria de la fuerza débil descubierta en 1983. Su masa ha sido medida por diversos experimentos, con distintos niveles de precisión. Los datos de la edición 2000 del Review of Particle Physics son, en GeV:
   Masa W - Experimento
   80.482 $\pm$ 0.091 - D0
   80.423 $\pm$ 0.112 - Aleph
   80.38 $\pm$ 0.12 - Opal
   80.61 $\pm$ 0.15 - L3
   80.41 $\pm$ 0.18 - CDF

   Muestre que el resultado combinado es 80.458 y analice la consistencia de los resultados con un test. Para esto asuma que las mediciones con incertezas gaussianas y construya el estadístico $S = \sum_i [(y_i - f(x_i))/\sigma_i]^2$, cuya distribución es χ^2 y calcule el *p-valor* del resultado.

   Rta: p=0.79

4. Dos grupos A y B formados cada uno de N=100 individuos, padecen una enfermedad. Se administra un suero al grupo A, pero no al grupo B (grupo control), siendo en todo lo demás los dos grupos tratados idénticamente. Se encuentra que en los grupos A y B, $n_1 = 75$ y $n_2 = 65$ individuos, respectivamente, se han recuperado de la enfermedad. Ensayar la hipótesis de que el suero ayuda a curar la enfermedad al nivel de significacia del (a) 0.01 (b) 0.05 (c) 0.10 . Para esto denote por $\hat{p_1}$ y $\hat{p_2}$ respectivamente las proporciones poblacionales curadas (1) utilizando el suero y (2) sin utilizar suero. Se debe decidir entonces entre las dos hipótesis:
   H_0: $p_1 = p_2$, las diferencias observadas son debidas al azar, es decir, el suero no es efectivo.
   H_1: $p_1 > p_2$, el suero es efectivo.
   Para realizar el test se decide construir el estadístico de prueba $z = (\hat{p_1} - \hat{p_2})/\sqrt{\sigma_{p_1}^2 + \sigma_{p_2}^2}$.
   Asumiendo que z sigue una distribución N(0,1) encuentre entonces el resultado esperado.

   Rta: (a) no se rechaza (se acepta) H_0 (b) no se rechaza (se acepta) H_0 (c) se rechaza H_0 (y se decide que el suero es efectivo)

Implementación de Herramientas Computacionales

1. Observe los valores densidad de probabilidad acumulativa y los valores-p para las distintas distribuciones presentadas en g5_cdf_ppf_values.py.

2. Utilice el programa g5_poisson_ccf.py para la construcción de cinturón de confianza para el caso de una variable con distribución de Poisson.

3. Visualice los discriminadores lineales y cuadráticos con el programa g5_plot_lda_qda.py

4. Tu primera Red Neuronal de Clasificación Básica *(Just for fun!)*: Predecir una imagen de moda con g5_tf_keras_classification.py .

5. Determine si ha cambiado la distribución de salarios en un dado país entre los años 1989 y 1990 mediante un test de Kolmogorov en base a g5_test_K_S.py