Handout 10

Pure significance tests

1. Let $X \sim N(m, \sigma^2 = 4)$. We want to test $H_0 : m = 0$ versus a general alternative $H_a : m \neq 0$, using a sample of $n = 100$ independent observations. We use $\overline{X}$ as a test statistic and find a value $\overline{x} = 0.384$
   (a) What would we conclude if we perform the significance test at the $\alpha = 0.05$ level?
   (b) What is the smallest significance level α at which the hypothesis would be rejected? (Note: such smallest α which still lead to the rejection of H_0 for a given value of the test statistic is called the p-value; the smallest p is, the stronger the evidence against H_0.)

2. A product is sold in packages, whose labeling states: "Net weight: 300g". Packaging is done by a machine which, when properly adjusted, produces packages with weight normally distributed, $N(300; \sigma^2 = 25)$.

   In order to test whether the machine is working properly, samples of 25 packages are taken periodically, their average weight $\overline{X}$ computed, and the process stopped if $\overline{X} < 300 - k$ or $\overline{X} > 300 + k$.
   (a) Find k such that when the machine is in fact producing packages with weight distributed $N(300; \sigma^2 = 25)$, the probability of stopping the process is 0.001.
   (b) For k fixed as in the preceding paragraph, find the probability of *not* stopping the process when it is yielding packages of mean weight 290g and variance $\sigma^2 = 25$.
   (c) Assume the process produces packages with mean m and variance $\sigma^2 = 25$. A sample of $n = 100$ packages gives an average weight $\overline{x} = 302.5$g. Give a 95% confidence interval for m.

3. To test that the mean λ of a Poisson distribution has value 1 we take a sample of $n = 9$ observations and obtain $\overline{X} = 1.55556$ What would be our conclusion at the $\alpha = 0.05$ significance level?

Respuestas abreviadas

1. (a) $\overline{X} \sim N(0, \sigma^2 = 0.04)$ bajo H_0. La región crítica sería bilateral, formada por dos colas de probabilidad 0.025. El valor que deja a su derecha una probabilidad de 0.025 en una $N(0.1)$ es:

```
> qnorm(0.975)
```

```
[1] 1.959964
```

Por tanto, la cola de la derecha comienza en

```
> 0 + 1.959964 * sqrt(0.04)
```

```
[1] 0.3919928
```

y la de la izquierda es simétrica. La región crítica es por tanto: $[0.392, \infty) \cup (-\infty, -0.392]$ (ó $[-0.392, 0.392]^c$). Al nivel de significación indicado, el valor 0.384 no estaría en la región crítica y no daría lugar al rechazo de H_0.

(b) Para que la hipótesis fuera rechazada, la región crítica debería comenzar en 0.384; la probabilidad conjunta de ambas colas sería entonces:

```
> 1 - ( pnorm(0.384/sqrt(0.04)) - pnorm(-0.384/sqrt(0.04)) )
```

```
[1] 0.0548579
```

2. (a) Requerimos k tal que $Prob(\overline{X} > 300 + k) = 0.001/2$. El valor dejnado a la derecha una probabilidad de 0.0005 en una $N(0, 1)$ es:

```
> qnorm(0.9995)
```

```
[1] 3.290527
```

Por tanto, k será:

```
> (k <- 3.2905 * sqrt(25/25) )
```

```
[1] 3.2905
```

y detendremos el proceso si $\overline{X}$ sale del intervalo $300 \pm 3.2905 = [296.7095, 303.2905]$.

(b) La probabilidad solicitada (que sería la de error de tipo II ó β) sería:

```
> pnorm( (303.2905-290) / 1 ) - pnorm( (296.7095-290) / 1 )
```

```
[1] 9.764634e-12
```

Tendríamos la práctica certeza de detectar una desviación tan grande como 10 gramos ($= 300 - 290$) en el peso medio de los paquetes envasados.

(c) Si la media es m tendremos que:

$$\frac{\overline{X} - m}{\sqrt{25/100}} \sim N(0,1)$$

y por tanto:

$$P\left(-1.96 \leq \frac{\overline{X} - m}{\sqrt{25/100}} \leq 1.96\right) = 0.95$$

o:

$$P\left(-1.96\sqrt{0.25} \leq \overline{X} - m \leq 1.96\sqrt{0.25}\right) = 0.95.$$

Operando:

$$P\left(\overline{X} - 1.96\sqrt{0.25} \leq m \leq \overline{X} + 1.96\sqrt{0.25}\right) = 0.95.$$

Reemplazando $\overline{X}$ por el valor obtenido en la muestra (302.5) obtenemos el intervalo de confianza $302.5 \pm 1.96\sqrt{0.25} = [301, 52, 303.48]$.

3. Al ser la variable discreta, no es posible construir una región crítica $\alpha = 0.05$ de modo exacto[1]. $\overline{X}$ no tiene distribución manejable, pero $S = n\overline{X}$ es $\mathcal{P}(\lambda = 9)$. En nuestro caso, toma el valor $9 \times 1.55556 \approx 14$. Si examinamos la distribución de una variable $\mathcal{P}(\lambda = 9)$,

```
> Fx <- ppois(0:20,lambda=9)
> names(Fx) <- 0:20
> Fx

           0            1            2            3
0.0001234098 0.0012340980 0.0062321951 0.0212264863
           4            5            6            7
0.0549636415 0.1156905208 0.2067808399 0.3238969643
           8            9           10           11
0.4556526043 0.5874082443 0.7059883203 0.8030083825
          12           13           14           15
0.8757734292 0.9261492307 0.9585336745 0.9779643408
```

[1]Se puede hacer, pero no de modo que veamos en el curso.

```
          16           17           18           19
0.9888940906 0.9946804287 0.9975735978 0.9989440463
          20
0.9995607481
```

vemos que colas totalizando probabilidad aproximada $\alpha = 0.05$ serían $[0, 3]$ y $[16, \infty)$ con probabilidad conjunta 0.021226 + (1 - 0.977964) = 0.0432. Incluso aunque incluyéramos 15 en la región crítica (lo que nos daría un nivel de significación 0.021226 + (1 - 0.958533) = 0.062693), el valor observado de S no estaría en la región crítica, y por lo tanto no rechazaríamos la hipótesis nula.

Orientaciones docentes

1. El apartado (1b) introduce un concepto nuevo, el de p-value o nivel de significación empírico: la probabilidad bajo la hipótesis nula de encontrar valores del estadístico igual o más raros que el encontrado. Deteneos un poco en ello, lo usarán constantemente en Econometría.

2. Haced notar en el apartado (2c) que no puede decirse del intervalo construido que contiene la media m con probabilidad 0.95 (la contiene, o no la contiene, pero no "la contiene con probabilidad 0.95"). Lo que puede decirse es que intervalos construidos mediante el mismo procedimiento con diferentes muestras, cubrirían la media con probabilidad 0.95.

   Conviene detenerse también en esto, porque da lugar a errores frecuentes. El *método* de construcción de intervalos de confianza $1 - \alpha$ tiene *probabilidad de cobertura* (confianza) de $1 - \alpha$. Por el contrario, *un* intervalo concreto construido mediante tal método tiene *probabilidad de captura* variable: ¡depende del intervalo! Puede ser mayor, igual o menor a $1 - \alpha$.