Handout 11

Testing against an alternative

1. In a previous problem we tested whether we were packing a product in packages with weight distributed as $N(300, \sigma^2 = 25)$. Assume that we do not care whether we are packing slightly more product per package, but we want to stop the process if the mean weight drops below 300g Find the most powerful critical region of size 0.001 when using a sample size $n = 100$.

2. We have $X \sim U(0, \theta)$, from which we take a random sample of size 2, (X_1, X_2). We want to test $H_0 : \theta \geq 2$ against the alternative $H_a : \theta < 2$. We adopt the rule of rejecting H_0 if $X_1 + X_2 \leq 0.5$

   (a) Argue on intuitive grounds whether said critical region is of the "right" form.

   (b) Find the significance level of the test.

   (c) Find the power when in fact $\theta = 1$.

3. In a dangerous crossroads, the interval of time between traffic accidents is exponentially distributed with mean 50 days. After a recent accident involving fatalities, local authorities reacted installing traffic lights, after which they claim the mean time between accidents will be 200 days.

   Residents in the area remain skeptical about the effectiveness of the measure, and the following statistical test is proposed: the hypothesis $H_0 : m = 50$ will be rejected in favor of the alternative $H_a : m = 200$ if the interval of time until the next accident is at least 150 days.

   (a) Compute the significance level of the test.

   (b) Compute the power of the test.

4. An optimal test is desired for $H_0 : X \sim N(0, 1)$ versus $H_a : X \sim N(1, 1)$. We want $\alpha = \beta = 0.05$. Find the form of the test and the sample size we need.

5. We want to test that the proportion p of defectives in an industrial process is $p = 0.02$ versus the alternative $p = 0.05$. Establish the decision rule if we want $\alpha = 0.05$ and we can afford a sample size $n = 100$.

6. Use Neyman-Pearson's theorem to find the most powerful critical region for the test of $H_0 : X \sim \mathcal{P}(\lambda = 0.5)$ versus $H_a : X \sim \mathcal{P}(\lambda = 1)$, such that $\alpha \leq 0.05$ and $n = 10$.

Respuestas abreviadas

1. Un estadístico adecuado sería $\overline{X}$ (suficiente para m en el caso de la normal, obtenible también como estadístico de contraste mediante el teorema de Neyman-Pearson). Claramente, la región crítica estará situada a la izquierda: son valores pequeños de peso promedio los que habremos de considerar evidencia contra H_0 en favor de la alternativa $H_a : m < 300$. Tenemos que encontrar un valor k tal que, bajo H_0 $P(\overline{X} < k) = 0.001$. Como bajo H_0 $\overline{X} \sim N(300, \sigma^2 = 0.25)$, y

   ```
   > qnorm(0.001)
   ```

   ```
   [1] -3.090232
   ```

   el valor de k deseado es $300 - 3.099232\sqrt{0.25}$

   ```
   > k <- 300 - 3.090232 * sqrt(0.25)
   > k
   ```

   ```
   [1] 298.4549
   ```

   Podemos en efecto comprobar que:

   ```
   > pnorm(k,m=300,sd=sqrt(0.25))
   ```

   ```
   [1] 0.001000001
   ```

2. El espacio muestral es un cuadrado $[0, \theta] \times [0, \theta]$ en el que cada muestra formada por dos valores es un punto (X_1, X_2).

   (a) La forma de la región crítica no parece insensata: rechazamos H_0 (θ "grande") si la muestra obtenida es un punto en las cercanías del origen, donde es menor la probabilidad de estar cuando mayor sea θ.

   (b) El nivel de significación es la probabilidad de que $X_1 + X_2 < 0.5$ cuando H_0 es cierta. Cuando, como ocurre aquí, H_0 es compuesta (no especifica un único valor de θ) el nivel de significación se calcula como el mayor α compatible con θ como prescribe H_0.

   En este caso, $\theta = 2$ es el que da mayor probabilidad a la región crítica. El cuadrado que acoge a todas las posibles muestras tiene área 4, la región crítica tiene área $\frac{0.5 \times 0.5}{2} = 0.125$ y siendo la distribución uniforme dentro del cuadrado, la probabilidad de la región crítica es $0.125/4 = 0.03125$.

   (c) Cuando $\theta = 1$, la potencia de la prueba es la probabilidad de estar en la región crítica, que tendría una probabilidad de $0.125/1 = 0.125$.

3. (a) El problema se reduce a contrastar $H_0 : \lambda = 1/50 = 0.02$ frente a $H_a : \lambda = 1/200 = 0.005$ con una única observación. La región crítica más potente estará formada por valores X verificando: verosimilitudes es:

$$\frac{0.005e^{-0.005X}}{0.02e^{-0.02X}} \geq k_\alpha$$

para k_α dependiente del nivel de significación. Tomando logaritmos, la región crítica más potente es de la forma:

$$\log(0.005) - 0.005X - \log(0.02) + 0.02X \geq \log(k_\alpha)$$

o equivalentemente,

$$X \geq \frac{\log(k_\alpha) - \log(0.005) + \log(0.02)}{0.015} = k'$$

La forma de decidir propuesta es un contraste más potente con $k' = 150$. El nivel de significación es la probabilidad de que, siendo H_0 cierta (es decir, siendo totalmente inefectiva la instalación de los semáforos), el intervalo X hasta el siguiente accidente exceda de 150. Como bajo H_0 la distribución es exponencial con $\lambda = 0.02$, dicha probabilidad es: $1 - F_X(150) = 1 - (1 - e^{-0.02\times150}) = 0.0497$.

(b) La potencia del contraste es la probabilidad de que $X > 150$ cuando $\lambda = 0.005$:

```
> 1 - (1 - exp(-0.005*150))
[1] 0.4723666
> 1 - pexp(150, rate=1/200)
[1] 0.4723666
```

4. Claramente, la región crítica será $[k, \infty)$ (se puede alternativamente obtener mediante Neyman-Pearson, si alguien no lo ve). Para que $\alpha = 0.05$,

$$1 - \Phi\left(\frac{k-0}{1/\sqrt{n}}\right) = 0.05$$

Para que $\beta = 0.05$,

$$\Phi\left(\frac{k-1}{1/\sqrt{n}}\right) = 0.05$$

De la primera igualdad despejamos:

$$\Phi\left(\frac{k-0}{1/\sqrt{n}}\right) = 0.95$$

y por tanto

$$\left(\frac{k-0}{1/\sqrt{n}}\right) = \Phi^{-1}(0.95) = 1.644854$$

Por lo tanto, sabemos que $k = 1.644854/\sqrt{n}$. Reemplazando este valor en la ecuación que da la condición $\beta = 0.05$ tenemos:

$$
\begin{aligned}
\Phi\left(\frac{1.644854/\sqrt{n} - 1}{1/\sqrt{n}}\right) &= 0.05 \\
\left(\frac{1.644854/\sqrt{n} - 1}{1/\sqrt{n}}\right) &= -1.644854 \\
1.644854/\sqrt{n} - 1 &= -1.644854/\sqrt{n} \\
3.289708/\sqrt{n} &= 1
\end{aligned}
$$

que proporciona $n = 10.82218$; como debe ser entero redondearemos a $n = 11$. El valor de k por su parte será: $1.644854/\sqrt{11} = 0.4959421$. Es fácil comprobar que con este valor de k, α y β son 0.05 o ligeramente menores:

```
> 1 - pnorm((0.4959421 -0)/ (1/sqrt(11)))      # alpha

[1] 0.04999998

> pnorm( (0.4959421 - 1) / (1/sqrt(11)))       # beta

[1] 0.04728475
```

5. El sentido común (o el teorema de Neyman-Pearson) hacen ver que la región crítica habrá de incluir muestras tales que $\overline{X}_n > k$, o alternativamente $\sum_{i=1}^{100} X_i > k'$. No se dice, pero el contexto sugiere que la muestra de $n = 100$ será una pequeña fracción de la población muestreada, y bajo este supuesto podemos suponer que $\sum_{i=1}^{100} X_i$ es aproximadamente $b(p = 0.02, n = 100)$ bajo la hipótesis H_0. A su vez, dicha distribución binomial es bien aproximable por una $\mathcal{P}(\lambda = 2)$. Buscamos pues un k' que deje a su derecha una probabilidad no mayor de $\alpha = 0.05$. Observando,

```
> ppois(0:8, lambda=2)

[1] 0.1353353 0.4060058 0.6766764 0.8571235 0.9473470 0.9834364 0.9954662
[8] 0.9989033 0.9997626
```

vemos que un tal valor es 4. La región crítica sera $[5, \infty)$ con un α de

```
> 1 - ppois(4, lambda=2)

[1] 0.05265302
```

Alternativamente podríamos encontrar k' como:

```
> qpois(0.95, lambda=2)

[1] 5
```

Si adoptamos como estadístico de contraste $\overline{X}_n$ la región crítica comenzará en 0.05.

6. Muy similar a un ejemplo realizado en clase, problema de aplicación de N-P completamente standard.

Orientaciones docentes

1. El ejercicio 2 tiene la ventaja de ilustrar nivel de significación y potencia en un caso desligado de las distribuciones usuales, y que obliga a pensar a partir de primeros principios.

2. El problema 4 es un poco diferente a los restantes en el sentido de que lo que se busca es el tamaño muestral necesario. Entronca bien así con cosas que veremos más tarde (en muestreo).

3. Los restantes problemas son completamente standard, intercambiables con cualesquiera otros.