Universidad Politécnica de Cartagena
Dpto. Matemática Aplicada y Estadística
Estadística *Variables aleatorias*

Problema 1

La dimensión de ciertas piezas sigue una distribución normal de media 150 y desviación típica 0.4. Sabiendo que se consideran aceptables todas aquellas piezas cuya longitud se encuentre dentro del intervalo (149'2 , 150'4). Determinar:

1. El porcentaje de piezas defectuosas.

2. Supongamos que se empaquetan en paquetes de 12 unidades, y un lote se rechaza si contiene más de 3 defectuosas. Determinar la proporción de lotes que se rechazarán.

3. Un determinado comprador decide comprarlas a granel en cajas de 360 unidades, pero no aceptará aquellas cajas con mas de 90 defectuosas. ¿Qué probabilidad tenemos de que nos acepte las cajas?. Comentar los resultados obtenidos en los dos últimos apartados.

Soluciones del Problema 1

Introducimos la v.a $X=$*"Dimensión de una pieza escogida al azar." Tenemos* $X \sim \mathcal{N}(150, 0.4^2)$.

1. *El intervalo (149'2 , 150'4) corresponde a* $(\mu-2\sigma, \mu+2\sigma)$ *y contiene por lo tanto el 95% de los valores de la distribución, lo que nos permite deducir que el porcentaje de piezas defectuosas es 5%.*

2. *El experimento es ahora "Rellenar un paquete con 12 piezas escogidas al azar entre la producción". Introducimos la v.a* $X=$*"número de piezas defectuosas en el paquete". Reconocemos la distribución Binomial, con parámetros* $n=12$*, y* $p=\mathbb{P}($*"una pieza escogida al azar es defectuosa"*$)$*. En el apartado anterior, hemos calculado* $p=0.05$*. Nos piden* $\mathbb{P}(X>3)$*. Usamos R Commander, Distribuciones, Distribuciones discretas, Distribución binomial, Probabilidades binomiales acumuladas, rellenamos*

los campos: valores de la variable 4 (empezamos a contar desde 4, puesto que nos interesa $X > 3$), número de ensayos: 12, probabilidad de éxito: 0.05, y queremos calcular la cola de la derecha. Obtenemos:

$$\mathbb{P}(X > 3) = 0.0002,$$

es decir que se recharián aprox. 2 cajas de cada 1000.

3. *El experimento es ahora: "Rellenar una caja con 360 piezas escogidas al azar entre la producción", introducimos la v.a Y="número de piezas defectuosas en la caja de 360 unidades.". Queremos calcular $\mathbb{P}(Y > 90)$. Y sigue una distribución binomial con parámetro $n = 360$, y parámetro $p = 0.05$ (al igual que en el apartado anterior). Usamos R Commander Distribuciones discretas, Distribución binomial, Probabilidades binomiales acumuladas, rellenamos los campos: valores de la variable 91 (empezamos a contar desde 91, puesto que nos interesa $X > 90$), número de ensayos: 360, probabilidad de éxito: 0.05, y queremos calcular la cola de la derecha. Obtenemos:*

$$\mathbb{P}(Y > 90) = 9e - 39,$$

que es muchísima más pequeña que en el apartado anterior a pesar de que 90 es la cuarta parte de 360, al igual que 3 es la cuarta parte de 12...

Problema 2

X es una variable aleatoria que sigue una distribución normal de media $\mu = 4$ y varianza desconocida . Sabiendo que el 58% de los valores de X se encuentran entre 3.25 y 4.75, calcular la varianza de X.

Soluciones del Problema 2

Si el 58% de los valores se encuentran en el intervalo centrado en $\mu = 4$, con extremos 3.25 y 4.75. Deducimos que el cuantil 0.21 es 3.25. Nos queda encontrar el valor de σ asociado. Una manera de hacerlo es ir tanteando con RCommander, cambiando el valor de σ. Usamos Distribuciones, Distribuciones continuas, Distribución normal, cuantiles de la distribución normal. La media es 4. La probabilidad es 0.21, y probamos con σ igual a 1 por ejemplo. Encontramos que el cuantil 0.21 es igual a 3.19, que es inferior al 3.25 objetivo. Tenemos por lo tanto que bajar algo la desviación típica. Probamos con $\sigma = 0.8$ y obtenemos un cuantil igual a 3.35, que es superior al 3.25 objetivo. Tanteando acabamos encontrando que $\sigma = 0.93$ nos proporciona el cuantil deseado.

Problema 3

El tiempo de duración de un ensamble mecánico en una prueba de vibración sigue una distribución exponencial de media 400 horas. Entonces:

1. Determinar la probabilidad de que el ensamble falle durante la prueba antes de 100 horas. ¿Cuál es la probabilidad de se produzca el fallo después de 500 horas?.

2. Si durante el ensayo se han probado 10 ensambles de manera independiente, determinar la probabilidad de que falle al menos uno de ellos antes de 500 horas. ¿Cuál sería la probabilidad de que fallasen todos transcurridas 800 horas?.

Soluciones del Problema 3

Introducimos la v.a. "Duración del ensamble en la prueba de vibración". Sabemos que $D \sim \mathcal{E}xp(\lambda)$*, y que* $\mathbb{E}[D] = 400$*. Vimos en clase que para una distribución exponencial,* $\mathbb{E}[D] = 1/\lambda$*. Deducimos por lo tanto que* $\lambda = 1/400$*.*

1. *Nos piden por una parte,* $\mathbb{P}(D < 100)$*. Usamos R commander, calculamos la probabilidad de la cola de la izquierda para una exponencial con parámetro 0.0025, y obtenemos*
$$\mathbb{P}(D < 100) = 0.22.$$
Por otra parte, debemos calcular $\mathbb{P}(D > 500)$*. Usando R Commander*
$$\mathbb{P}(D > 500) = 0.29.$$

2. *El experimento es ahora: "Probar 10 ensambles escogidos al azar", y la variable que nos interesa primero es N1="Número de ensambles que han fallado antes de 500 horas". Nos piden* $\mathbb{P}(N1 \geq 1)$*. La distribución de* $N1$ *es una Binomial con parámetros* $n = 10$ *y* $p = \mathbb{P}($*"un ensamble escogido al azar falla antes de 500 horas*$)$*. Hemos calculado en el apartado anterior la probabilidad complementaria, y deducimos* $p = 0.71$*. Por otra parte, usamos* $\mathbb{P}(N1 \geq 1) = 1 - \mathbb{P}(N1 = 0)$*. Calculamos* $\mathbb{P}(N1 = 0)$ *usando RCommander, Distribuciones, Distribuciones discretas, Distribución binomial, probabilidades acumuladas de una binomial. Encontramos* $\mathbb{P}(N1 \geq 1) = 0.99989$*.*

Para contestar a la pregunta "Cuál sería la probabilidad de que fallasen todos transcurridas 800 horas?", introducimos, para el mismo experimento aleatorio la variable N2="Número de ensambles que han fallado antes de 800 horas". También se trata de una distribución Binomial. Sus parámetros son $n = 10$, *y* p, *la probabilidad de que un ensamble escogido al azar falle antes de 800 horas. Es decir que* $p = \mathbb{P}(D \leq 800)$. *Usamos R Commander como en el primer apartado para encontrar:* $p = 0.86$. *Queremos calcular* $\mathbb{P}(N2 = 10)$. *Usamos R Commander, al igual que para el párrafo anterior y encontramos:* $\mathbb{P}(N2 = 10) = 0.22$.

Problema 4

El número de visitas realizadas en un día entre semana en una determinada página web se decide modelizar por una variable de Poisson de media 8. Se pide:

1. ¿Cuál es la probabilidad de que en un día se reciban más de 4 visitas?

2. Y ¿entre 7 y 10 visitas (ambos incluidos)?

Soluciones del Problema 4

Introducimos la v.a V *="Número de visitas a la página web en un día entre semana escogido al azar". Sabemos que* $V \sim \mathcal{P}(\lambda)$, *y por otra parte* $\mathbb{E}[V] = 8$. *Vimos en clase que la media de una distribución de Poisson es* λ. *Por lo tanto* $V \sim \mathcal{P}(8)$.

1. *Nos piden* $\mathbb{P}(V > 4)$. *Usamos R Commander, en el menú Distribuciones, Distribuciones discretas, Poisson, Probabilidades de Poisson acumuladas. Introducimos el valor de la variable 5 (porque V debe ser estrictamente mayor de 4), la media igual a 8, escogemos la cola de la derecha y encontramos 0.81.*

2. *Nos piden* $\mathbb{P}(7 \leq V \leq 10)$. *Usamos la diferencia entre dos probabilidades acumuladas:*

$$\mathbb{P}(7 \leq V \leq 10) = \mathbb{P}(V \leq 10) - \mathbb{P}(V \leq 6)$$

Usamos R Commander, en el menú Distribuciones, Distribuciones discretas, Poisson, Probabilidades de Poisson acumuladas. Introducimos dos valores separados por espacio: 6 10, la media igual a 8, escogemos la cola de la izquierda, obtenemos

$$\mathbb{P}(V \leq 10) = 0,82, \qquad \mathbb{P}(V \leq 6) = 0.31.$$

Deducimos

$$\mathbb{P}(7 \leq V \leq 10) = 0.51.$$

Problema 5

Decidimos modelizar el valor proporcionado por un aparato de medición de peso, por una variable aleatoria X con distribución Normal de media μ y desviación típica $\sigma = 0.5$.

1. ¿Qué quiere decir que la variable X es *aleatoria*? En su opinión, ¿cuál es el procedimiento que nos ha llevado a escoger un modelo Normal para la distribución de los valores de X? ¿Qué representa la cantidad μ? ¿Cuándo decimos que un aparato es preciso? ¿y exacto? [**0.75pt/10**]

2. Supongamos que $\mu = 13$, ¿cuál es la probabilidad de que el valor medido sea mayor de 14? ¿y menor que 12? [**0.75pt/10**]

3. Suponiendo que $\mu = 13$, proporcionar un intervalo centrado en 13 que contenga el 95% de los valores medidos. [**0.75pt/10**]

4. Decidimos repetir 4 veces la medición y calcular la media de las cuatro mediciones. ¿Cuál es la distribución de los valores obtenidos? ¿Por qué es mejor proporcionar la media de 4 mediciones en lugar de sólo medir una vez? [**0.75pt/10**]

Soluciones del Problema 5

1. *El decir que una variable es aleatoria quiere decir que es imposible predecir su valor, sólo podemos especificar qué con probabilidad pensamos que tome un determinado valor o pertenezca a un determinado intervalo. El escoger un modelo Normal para la distribución de valores de la variable X, se basa en examinar un histograma de bastante valores medidos para ver si se ajusta a una campana de Gauss. La cantidad μ representa el centro de la distribución de los valores proporcionados por el aparato de medición. Si el aparato de medición es exacto, este centro (μ) debe coincidir con el valor exacto de la cantidad que buscamos medir. El aparato es preciso si la desviación típica de los valores proporcionados es pequeña, es decir que la distribución de los valores medidos presenta poca dispersión. (ver tema 5)*

2. *Tenemos que* $X \sim \mathcal{N}(\mu, \sigma^2)$, *con* $\mu = 13$ *y* $\sigma = 0.5$. *Para calcular una probabilidad asociada a* X *basta con tipificar:*
Por ejemplo,
$$\mathbb{P}(X > 14) = \mathbb{P}(\frac{X-13}{\sigma} > \frac{14-13}{\sigma}) = \mathbb{P}(Z > 2) = 1 - \phi(2) = 0.0228.$$
En cuanto a $\mathbb{P}(X < 12)$, *utilizamos la simetría de la campana de Gauss respecto a su media para deducir que* $\mathbb{P}(X < 12) = \mathbb{P}(X > 14)$.

3. *Sabemos que, para una variable Normal, el 95% de los valores se encuentran entre* $\mu - 2\sigma$ *y* $\mu + 2\sigma$. *Por lo tanto el intervalo que buscamos es* $13 - 2 \times 0.5 = 12$ *y* $13 + 2 \times 0.5 = 14$

4. *Sea Sean* X_1, X_2,... X_4, *las variables "valor obtenido en la 1º médición", etc hasta "valor obtenido en la 4º medición". Consideramos la media muestral* $\bar{X} = \frac{X_1+X_2+\cdots+X_4}{4}$.
Puesto que $X_1, \cdots, X_4$ *forman una muestra aleatoria simple de la distribución* $\mathcal{N}(\mu, \sigma^2)$, *sabemos*
$$\bar{X} \sim \mathcal{N}(\mu, \sigma^2/4),$$
es decir $\bar{X} \sim \mathcal{N}(\mu, 0.0625)$. *La desviación típica de los valores de* $\bar{X}$ *es dos veces más pequeña que la de* X. *Al promediar mediciones obtenemos por lo tanto una mayor precisión.*

Problema 6

De un determinado prefabricado de escayola se estudian dos características, su "grosor" y su "longitud". Se sabe que ambas características siguen una distribución Normal y que la variable longitud tiene media 200 cm y desviación típica 0.5 cm, mientras que la variable grosor tiene media y desviación típica desconocidas. Sin embargo, se sabe por mediciones anteriores el 50% de los valores de grosor son mayores que 15mn, y que el 2.5% de las piezas presentan un grosor inferior a 14 mm.

1. Determinar la media y la desviación típica de la variable grosor.

2. Para que una pieza pueda ser aprovechada, debe tener una longitud comprendida entre 199 cm y 201 cm. Determinar el porcentaje de piezas que no cumplen las especificaciones de longitud.

3. Supongamos que los prefabricados se empaquetan en lotes de 500 unidades, determinar la probabilidad de que un paquete contenga menos de 15 piezas que no cumplen la norma de longitud.

Soluciones del Problema 6

Introducimos las variables G y L que son el grosor y la longitud de una pieza escogia al azar. Sabemos que $L \sim \mathcal{N}(200, 0.5^2)$, y que G también sigue una distribución normal.

1. *Nos dicen que la mediana de G es 14, y que el percentil 2.5 es igual a 15mn. Para una normal, la mediana coincide con la media, por lo tanto $\mu_G = 15$. Además puesto que, para una normal, el intervalo $(\mu - 2\sigma, \mu + 2\sigma)$ contiene el 95% de los valores, el percentil 2.5% es igual a $\mu - 2\sigma$. Deducimos que la desviación típica de G también es 0.5mm.*

2. *Para la variable L, el intervalo $(199, 201)$ corresponde a $(\mu - 2\sigma, \mu + 2\sigma)$, por lo tanto, el porcentaje de piezas que no cumplen las especificaciones es 5%.*

3. *El experimento es ahora "escoger al azar 500 unidades para rellenar un paquete". La v.a. X="Número de piezas que no cumplen la norma de la longitud". Es una variable con distribución binomial con parámetros $n = 500$ y p, la probabilidad de que, al escoger una unidad al azar me salgan que no cumple la norma de longitud. Hemos calculado en el apartado anterior $p = 0.05$. Tenemos que calcular $\mathbb{P}(X < 15)$. Usamos R Commander, Distribuciones, Distribuciones discretas, Distribución binomial, probabilidades acumuladas de una binomial. Rellenamos los campos: valores de la variable 14 (empezamos a contar desde 14 hacia abajo, puesto que nos interesa $X < 15$), número de ensayos: 500, probabilidad de éxito: 0.05, y queremos calcular la cola de la izquierda. Obtenemos:*

$$\mathbb{P}(X < 15) = 0.01.$$