Repaso de conceptos de teoría (ver bibliografía: Meyer)

Sea ε un experimento aleatorio y S el espacio muestral correspondiente, definimos una función X que asigna a cada uno de los elementos de $s \in$ S un único número real x=X(s). Llamaremos a X como *variable aleatoria*, y el conjunto de todos los valores posibles de X será el *recorrido* R_X.
Sea un suceso A ⊂ S y un suceso B ⊂ R_X, y si $\mathrm{A} = \{s \in \mathrm{S}/\mathrm{X}(s) \in \mathrm{B}\}$, decimos que A y B son *sucesos equivalentes*. Luego, como ya sabemos calcular la probabilidad P(A), y siendo que A y B son equivalentes, definimos a la probabilidad de B como P(B)=P(A).

Valores característicos
Esperanza matemática o valor esperado E(X)

$$\mathrm{E}(\mathrm{X}) = \sum_{\forall i} x_i \, \mathrm{P}(\mathrm{X} = x_i) = \sum_{\forall i} x_i \, \mathrm{p}(x_i)$$

Varianza V(X) y desvío estándar σ_X

$$\mathrm{V}(\mathrm{X}) = \mathrm{E}[(\mathrm{X} - \mathrm{E}(\mathrm{X}))^2] = \mathrm{E}(\mathrm{X}^2) - \left(\mathrm{E}(\mathrm{X})\right)^2 = \sigma_{\mathrm{X}}^2$$

Propiedades
Si C es una constante, y X e Y variables aleatorias cualesquiera:

$E(C)=C$
$E(C\,X)=C\,E(X)$
$E(X+Y)=E(X)+E(Y)$

$V(X+C)=V(X)$
$V(C\,X)=C^2\,V(X)$

Si $X_1, X_2, \dots, X_n$ son variables aleatorias *independientes*:

$V(X_1+X_2+\dots+X_n)=V(X_1)+V(X_2)+\dots+V(X_n)$

<u>Variables aleatorias discretas con distribuciones conocidas</u>

Experimento de Bernoulli: es un experimento aleatorio ε_B que tiene solamente dos resultados posibles, “éxito” y “no éxito”.
Si definimos el suceso D:”éxito”, la probabilidad de éxito será P(D)=p y la probabilidad del suceso complementario será $P(\overline{D})=1-p=q$

Sea la variable aleatoria discreta X definida como

X:"cantidad de éxitos en una única repetición del experimento $\mathcal{E}_B$"

Justificar

Definiendo el suceso D:"éxito", y siendo P(D)=*p* (y $P(\overline{D})=1-p=q$)

X tendrá una distribución de Bernoulli de *parámetro p* y de denota como
$X \sim B_e(p)$

Porque se puede tener 1 éxito o ninguno

El recorrido es $R_X=\{0,1\}$ o también puede expresarse como $x = 0{,}1$

y su función de probabilidad $P(X=x)=p_X(x)$ será

$P(X = x) = p_X(x) = p^x(1-p)^{1-x} = p^x\, q^{1-x}$ con $x = 0{,}1$

No olvidar indicar R_X

$E(X)=p$
$V(X)=p\, q=p\,(1-p)$

Justificar ambos resultados

En esta unidad se estudiarán varias distribuciones de probabilidad basándonos en este tipo de experimentos $\mathcal{E}_B$. Se sugiere completar la tabla disponible en el material de la unidad con la información de cada distribución estudiada.

En el desarrollo anterior se usaron dos notaciones posibles $P(X=x)$ y $p_X(x)$ para la función de probabilidad. Lógicamente, con usar una de las notaciones será suficiente. Por el momento, las notaciones $p_X(x)$ y $p(x)$ son equivalentes porque nos referimos a una sola variable aleatoria X. Prestar especial cuidado al uso de letras mayúsculas y minúsculas, *cursiva* e imprenta. Recordar además que hay algunas hipótesis que deben cumplirse para considerar que una dada variable aleatoria tiene una distribución determinada.

En el ejercicio 12 de la práctica se incluye una discusión resumida de procesos de Poisson. Se aconseja entender claramente la resolución del mismo antes de resolver el resto de los ejercicios referidos a procesos de Poisson (ejs. 10, 11 y 13 y adicionales).

Algunos ejercicios resueltos

<u>**Variables aleatorias discretas**</u>

<u>Generalidades</u>

3) La siguiente tabla muestra la función de distribución acumulada de una

variable aleatoria X

x	1	2	3	4
$P(X\leq x_i)=F(x_i)$	1/8	3/8	3/4	1

Determine:

a) la función de probabilidad puntual de X.

Calculamos la función de probabilidad puntual $P(X=x_i)=p(x_i)$ como

P(X=1)=p(1)=F(1)=1/8

p(2)=F(2)-F(1)=2/8=1/4

p(3)=F(3)-F(2)=3/8

p(4)=F(4)-F(3)=1/4

Podemos resumir los resultados en la tabla siguiente

x	1	2	3	4
$P(X=x_i)=p(x_i)$	1/8	1/4	3/8	1/4

b) P(1 ≤ X ≤ 3), P(X < 3) y P(X > 1,4).

Hay muchas formas de evaluar estas probabilidades

P(1 ≤ X ≤ 3)=P(X=1)+P(X=2)+P(X=3)=F(3)-F(1)+p(1)=…

P(X < 3)=P(X ≤ 2)=F(2)=…

P(X > 1,4)=p(2)+p(3)+p(4)=1-P(X ≤ 1)=…

c) Calcule e interprete la esperanza matemática, varianza y desvío estándar de X.

$$\mathrm{E(X)} = \sum_{i=1}^{4} x_i\ \mathrm{p}(x_i) = 1\cdot\frac{1}{8} + 2\cdot\frac{1}{4} + 3\cdot\frac{3}{8} + 4\cdot\frac{1}{4} = \frac{11}{4}$$

$$\mathrm{E(X^2)} = \sum_{i=1}^{4} x_i^2\ \mathrm{p}(x_i) = \frac{17}{2}$$

$$\mathrm{V(X)} = \mathrm{E(X^2)} - [\mathrm{E(X)}]^2 = \frac{15}{16} \qquad \sigma_{\mathrm{X}} = \sqrt{V(X)} = \cdots$$

E(X) representa …

σ_X representa la dispersión de los valores de X respecto de E(X).

Se debe resaltar que usamos la notación x_i para simbolizar un valor específico que puede asumir X (la variable aleatoria). Esta notación es necesaria cuando formulamos las ecuaciones para calcular E(X) y V(X). Vamos a ver que en el resto de la práctica usaremos también x para referirnos a un valor específico de X.

Algunas variables aleatorias discretas famosas

6) De un lote de 20 artículos 4 son defectuosos. Se eligen al azar 3 artículos. Sea X: "número de artículos defectuosos encontrados". Determine la distribución de X cuando los artículos se extraen con reposición y cuando se extraen sin reposición. En ambos casos calcular la función de probabilidad acumulada.

El problema nos pide analizar la cantidad de artículos defectuosos al hacer extracciones de un lote. Es decir, se hace una extracción y se analiza si el artículo es defectuoso o no, y luego se repite el experimento. El éxito del experimento es que el artículo extraído se a defectuoso.
La extracción con reposición significa que se devuelve el artículo extraído al lote cada vez que se repite el experimento.

Caso 1: extracción con reposición
Definimos suceso D:"artículo extraído es defectuoso"
P(D)=4/20=0,2=*p*
$P(\overline{D})$=0,8=*q*

Justificar

Definimos con mayor precisión la variable aleatoria
X:"número total de artículos defectuosos hallados al extraer 3 artículos al azar"

Suponiendo que cada extracción es independiente de las demás, podemos decir que X tiene una distribución Binomial de *parámetros* 3 y 1/5 y lo denotamos así X~Bi(3,1/5)

La función de probabilidad es

$$P(X = x) = \binom{3}{x} \left(\frac{1}{5}\right)^{x} \left(\frac{4}{5}\right)^{3-x} \quad \text{con } x = \overline{0,3}$$

La función de distribución acumulada F(x)=P($X \leq x$) resulta

$$F(x) = P(X \leq x) = \sum_{i=0}^{x} \binom{3}{i} \left(\frac{1}{5}\right)^{i} \left(\frac{4}{5}\right)^{3-i} \quad \text{con } x = \overline{0,3}$$

F(0)=p(0)=…
F(1)=p(0)+p(1)=…
F(2)=…
F(3)=…

Se puede resolver usando la tabla provista en el material de clase

El problema no lo pide, pero

$$\mathrm{E(X)} = 3 \cdot \frac{1}{5} = 0{,}6 \qquad \mathrm{V(X)} = 3 \cdot \frac{1}{5} \cdot \frac{4}{5} = 0{,}48$$

Caso 2: extracción sin reposición
En este caso, P(D) variará en cada repetición del experimento, dependiendo de si encontramos o no un elemento defectuoso. Además, la cantidad de piezas del lote es comparable a la cantidad de extracciones.

Definimos con mayor precisión la variable aleatoria
X:"número total de artículos defectuosos hallados al extraer 3 artículos al azar de un total de 20 artículos"

Suponiendo que cada extracción es independiente de las demás, podemos decir
X~Hi(20,4,3)
La función de probabilidad es

$$\mathrm{P(X} = x) = \frac{\binom{4}{x}\binom{20-4}{3-x}}{\binom{20}{3}} = \frac{\binom{4}{x}\binom{16}{3-x}}{\binom{20}{3}} \quad \text{con } x = \overline{0{,}3}$$

Nota: analizar P(X=*x*) en función de los conceptos del capítulo 3, es decir
-formas posibles de extraer 3 elementos de un total de 20 es $\binom{20}{3}$
-formas posibles de extraer *x* elementos defectuosos de un total de 4 defectuosos es …
- formas posibles de extraer 3-*x* elementos no defectuosos de un total de 16 no defectuosos es ….

La función de distribución acumulada F(*x*)=P($\mathrm{X} \leq x$) resulta

$$\mathrm{F}(x) = \sum_{i=0}^{x} \frac{\binom{4}{i}\binom{16}{3-i}}{\binom{20}{3}} \quad \text{con } x = \overline{0{,}3}$$

F(0)=…
F(1)=…
F(2)=…
F(3)=…

Se debe usar la fórmula de p(*x*) porque no se dispone de tablas

El problema no lo pide, pero

$$\mathrm{E(X)} = 3 \cdot \frac{4}{20} = 3 \cdot \frac{1}{5} = 0{,}6$$

$$\mathrm{V(X)} = 3 \cdot \frac{4}{20} \cdot \frac{(20-4)}{20} \cdot \frac{(20-3)}{(20-1)} = 3 \cdot \frac{1}{5} \cdot \frac{4}{5} \cdot \frac{17}{19} = \cdots$$

Relacionar E(X) y V(X) de este caso con el anterior de este ejercicio.

8) Supongamos que un mecanismo es inspeccionado regularmente al finalizar cada día para chequear si aún funciona correctamente. Sea *p* la probabilidad del suceso "el mecanismo falla durante cualquier día dado". Considere la variable aleatoria X: "número de inspecciones necesarias para obtener la primera falla".

Suponemos que el estado de funcionamiento del mecanismo al hacer la inspección es independiente del estado encontrado en todas las otras inspecciones.

Sea el suceso F: "el mecanismo falla durante un día dado"
P(F)=*p* y P($\overline{\mathrm{F}}$)=1-*p*=*q*
a) Determine la distribución de probabilidades de X.
X: "número de inspecciones necesarias para obtener la primera falla"
X~Ge(*p*)

$$\mathrm{P(X} = x) = p\, q^{x-1} \qquad \text{con } x = 1, 2, \cdots$$

También podemos expresar el recorrido como $\mathrm{R_X} = \mathbb{N}$

b) Halle la probabilidad de que sean necesarios 5 días para encontrar la primera falla.

$$\mathrm{P(X} = 5) = \mathrm{p}(5) = p\, q^{5-1} = p\, q^4$$

c) Halle *p* de modo que la probabilidad hallada en b) sea máxima.
Sabemos que p(5) es una función de *p*, por lo que el problema se reduce a encontrar el valor p_M que maximice la función *f*(*p*).
Se sugiere resolver el problema expresando p(5) como una función de *q*=1-*p*
Finalmente, resulta p_M=1/5

$$\text{Luego } \mathrm{P(X} = 5) = \frac{1}{5}\left(\frac{4}{5}\right)^4 \cong 0{,}082$$

El problema no lo pide, pero para el valor de *p* calculado

$\mathrm{E(X)} = 5$

$\mathrm{V(X)} = \frac{4}{5} \cdot 5^2 = 20$

9) Un "gran" lote de llantas para autos contiene el 20% de llantas defectuosas. De ese lote se seleccionarán 5 llantas para colocar en un auto.

Como disponemos de un "gran" lote de llantas, podemos decir que la proporción de llantas defectuosas se mantiene constante a pesar que el experimento sea sin reposición.
Sea el suceso D:"llanta defectuosa"
P(D)=0,2 y P($\overline{\mathrm{D}}$)=0,8

a) Calcule la probabilidad de que sea necesario seleccionar 8 llantas del lote para obtener 5 en buen estado.
Definimos la variable aleatoria
X:"número de llantas seleccionadas hasta obtener 5 llantas en buen estado"
X~Pa(5, 0,8)

$\mathrm{P(X} = x) = \binom{x-1}{5-1}\ 0{,}8^5\ 0{,}2^{x-5} \quad \text{con } x = 5, 6, \cdots$

P(X=8)=...

b) Calcule el número promedio de selecciones para obtener 5 llantas en buen estado.
Debemos calcular la esperanza

$\mathrm{E(X)} = \frac{5}{0{,}8} = 6{,}25$

Luego, el número promedio de selecciones que debemos realizar para obtener 5 llantas en buen estado es 6,25. Debemos recordar que E(X) puede no ser un valor posible de la v.a. discreta X, como en este caso.
Luego, si repetimos el experimento un gran número de veces y calculamos el promedio del número de llantas seleccionadas, veremos que ese promedio se acerca al valor de E(X).

El problema no lo pide, pero

$\mathrm{V(X)} = \frac{5 \cdot 0{,}2}{0{,}8^2} = \cdots$

12) El número de clientes que entran en un negocio es una variable aleatoria de Poisson con tasa 30 clientes/hora.

En este problema se contabiliza la cantidad de clientes que ingresan a un negocio en un intervalo de tiempo de duración *t*. Esta cantidad de clientes que ingresan en un intervalo *t* es aleatoria y tiene una distribución de Poisson (del enunciado).

El valor de la tasa de 30 clientes/hora significa que, en un período de tiempo de 1 hora se espera que ingresen 30 clientes en total. Si en cambio consideramos un intervalo de 2 horas, y si no cambió ninguna condición (por ejemplo, el local no cerró en ese intervalo de tiempo), se espera que ingresen 60 clientes en total. Si consideramos un intervalo de 20 minutos, la cantidad esperada de clientes será de 10.

En los tres casos mencionados, estamos hablando de tres variables aleatorias con distribuciones de Poisson diferentes porque el intervalo *t* considerado es distinto. Debemos notar que estas vv.aa. están relacionadas por la misma tasa de entrada de clientes por unidad de tiempo. Como la duración del intervalo *t* puede ser cualquiera, estamos hablando de un conjunto o colección de vv.aa. discretas, donde cada v.a. tiene asociado un período *t* particular.

Expresado de otra forma, esta colección está formada por vv.aa. discretas con distribuciones de Poisson porque el experimento sigue siendo el conteo de éxitos, pero las distribuciones son diferentes porque la magnitud continua *t* que es la que limita el experimento es distinta para cada v.a. particular. En estos casos, se dice que estamos hablando de un *proceso de Poisson*. La magnitud continua a la que nos referimos puede ser un intervalo tiempo, longitud, área, volumen u otra magnitud continua relacionada con el problema.

Para identificar a la colección de vv.aa. emplearemos la convención ya usada para identificar una v.a. (una letra mayúscula) agregando un subíndice que representa al intervalo, por ejemplo X_t. Si queremos referirnos a una v.a. particular que pertenece a esta colección, reemplazamos el subíndice *t* por el valor particular del intervalo para dicha v.a..

Como en el ejercicio analizamos intervalos en minutos, es decir, valores de *t* en minutos, expresaremos la tasa (o cantidad de éxitos por unidad de tiempo) en clientes por minuto. Convencionalmente identificaremos a esta tasa con λ.

Entonces, la tasa λ será

$$\lambda = 30\frac{clientes}{hora} = 30\frac{clientes}{hora}\cdot\frac{hora}{60\ minutos} = \frac{1}{2}\frac{clientes}{minuto}$$

Luego, identificaremos
X_t:"número de clientes que entran al negocio en *t* minutos"

$$X_t \backsim \mathrm{Po}\left(\frac{1}{2}t\right)$$

Notar que el parámetro usado es λt=1/2 *t*

$$P(X_t = x) = e^{-\frac{t}{2}}\frac{\left(\frac{t}{2}\right)^x}{x!} \quad \text{con } x = 0, 1, 2, \cdots$$

$E(X_t)=t/2$ $\qquad$ $V(X_t)=t/2$

a) Calcule la probabilidad de que en un período de 2 minutos no ingresen clientes al negocio.
En este caso *t*=2 minutos, y la v.a. es
X_2:"número de clientes que entran al negocio en 2 minutos"

$$P(X_2 = 0) = e^{-\frac{2}{2}} = e^{-1} \cong 0{,}3679$$

b) Se observa el número de clientes que ingresan al negocio durante 15 intervalos de 2 minutos. Calcule la probabilidad de que en a lo sumo 5 de esos intervalos no se registren ingresos.

Definimos el suceso C:"no ingresan clientes en un intervalo de 2 minutos"

$$P(C) = P(X_2 = 0) \cong 0{,}3679$$

Se puede calcular usando la tabla provista

Luego, definimos la v.a. discreta Z como
Z:"número de intervalos de 2 minutos en los que no ingresan clientes de un total de 15 intervalos de 2 minutos observados"
Z~Bi(15, 0,3679)

La función de probabilidad es

$$P(Z = z) = \binom{15}{z} 0{,}3679^z\, 0{,}6321^{15-z} \qquad \text{con } z = \overline{0{,}15}$$

La probabilidad de que en a lo sumo 5 de esos intervalos no se registren ingresos resulta

$$P(Z \leq 5) = \sum_{z=0}^{5} \binom{15}{z}\ 0{,}3679^{z}\ 0{,}6321^{15-z} = \cdots$$

El problema no lo pide, pero
E(Z)=... V(Z)=...

Ejercicios adicionales

5) En una fábrica dos máquinas producen el mismo artículo. La máquina A produce diariamente el doble de artículos que la máquina B. El 2% de los artículos producidos por la máquina A son defectuosos al igual que el 3% de los producidos por B. Si se combina la producción diaria de ambas máquinas y luego se toman al azar 10 artículos. ¿Cuál es la probabilidad de encontrar a lo sumo dos defectuosos?

Definimos una serie de variables aleatorias:
N:"cantidad total de artículos fabricados"
Na:"cantidad de artículos fabricados por la máquina A"
Nb:"cantidad de artículos fabricados por la máquina B"
R:"cantidad total de artículos defectuosos fabricados"
Ra:"cantidad total de artículos defectuosos fabricados por la máquina A"
Rb:"cantidad total de artículos defectuosos fabricados por la máquina B"

De los datos del problema, podemos calcular

Na=2 Nb (1)

De (1)

Ra=0,02 Na=0,02 . 2 Nb=0,04 Nb (2)

Rb=0,03 Nb (3)

De (1)

N=Na + Nb=2 Nb + Nb=3 Nb (4)

De (2) y (3)

R=Ra+Rb= 0,04 Nb + 0,03 Nb= 0,07 Nb (5)

Definimos el suceso D:"el artículo encontrado es defectuoso"

$$P(D) = \frac{R}{N} = \frac{0{,}07\ Nb}{3\ Nb} = \frac{7}{300}$$

De (4) y (5)

Consideramos que cada artículo fue fabricado independientemente del resto.
Entonces, analizaremos la v.a. discreta
X:"número de artículos defectuosos en un grupo de 10 artículos"

Debemos tener en cuenta dos casos.
Caso 1: la cantidad total N de artículos producidos es suficientemente grande como para que P(D) no varíe apreciablemente (sería como extracción con reposición). Entonces, X tendrá una distribución binomial.

Caso 2: la cantidad total N de artículos producidos no es suficientemente grande (sería como extracción sin reposición). Entonces X tendrá una distribución hipergeométrica.

Completar la resolución del ejercicio para ambos casos.

6) Un sistema de protección está constituido por *n* radares que funcionan independientemente. La probabilidad de que un radar detecte un cohete que ingrese a la zona de protección es 0,9. ¿Cuál es el número de radares necesarios para que la probabilidad de detectar un cohete que ingrese a la zona de protección sea 0, 999?

Definimos el suceso
D:"radar detecta un cohete"
P(D)=0,9

Suponemos que los radares funcionan independientemente unos de otros.
Si tenemos un conjunto de *n* radares, habrá detección si al menos un radar detecta un cohete. Entonces, definimos la v.a. discreta
X:"número de radares que detectan un cohete de un total de *n* radares"
X~Bi(*n*, 0,9)

$$\mathrm{P}(\mathrm{X}=x)=\binom{n}{x}\ 0{,}9^{x}\ 0{,}1^{n-x} \qquad \text{con } x=\overline{0,n}$$

Como dijimos antes, habrá detección si al menos un radar detecta un cohete.
Entonces, planteamos

$$\mathrm{P}(\mathrm{X}\geq 1)=1-\mathrm{P}(\mathrm{X}=0)=0{,}999$$

De la igualdad anterior, despejamos el valor de *n* que se pide.

9) El número de artículos demandados por semana en un negocio es una v.a. de Poisson con tasa 7 artículos/semana. ¿Cuántos artículos debe tener dicho negocio al comenzar la semana, para satisfacer la demanda con probabilidad mayor o igual que 0,99?

Definimos la colección de variables (proceso de Poisson)
X_t:"número de artículos demandados en *t* semanas"
$X_t \backsim \mathrm{Po}(7t)$
$$P(X_t = x) = e^{-7t}\frac{(7t)^x}{x!} \quad \text{con } x = 0, 1, 2, \cdots$$

Aunque el problema no lo pide: $E(X_t)=7t$ $V(X_t)=7t$

Como el problema pregunta el stock del artículo que debe tener el negocio al comenzar la semana para satisfacer la demanda, debemos analizar la v.a. X_1 porque t=1 semana.

Supongamos que el negocio tiene un stock de *k* artículos al comenzar la semana. Satisfacer la demanda con probabilidad mayor o igual que 0,99 significa que
$$P(X_1 \leq k) = \sum_{x=0}^{k} e^{-7}\frac{7^x}{x!} \geq 0{,}99$$

El problema se resuelve usando la tabla de la distribución de Poisson, sumando hasta que la probabilidad sea mayor o igual que 0,99. De esta manera obtenemos el valor de *k*.

Finalmente, al inicio de la semana el negocio debe tener un stock de … artículos para satisfacer la demanda con una probabilidad mayor o igual a 0,99.