Semana Tec: Estadística y probabilidad con juegos de mesa
Análisis sobre el mejor punto de inicio para colocar un aldea en Catan basado en la disposición que sugiere el juego

Descripción del juego o escenario a acotar (SING0202 Interpretación de variables)

Catán(compra de aldeas, cuidades e intercambio de recursos)
Es un juego de mesa que se basa en la idea de desarrollo de civilizaciones y colonización. En él se tienen de 3 a 4 jugadores, donde cada uno quiere conquistar a la isla de Catán, y existen aldeas, ciudades, carreteras y cartas de desarrollo. Todos los jugadores comienzan con dos aldeas y dos carreteras. El objetivo del juego es llegar a los 10 puntos de victoria, donde cada aldea vale 1 punto de victoria, y cada ciudad vale 2 puntos de victoria. Las cartas de desarrollo le dan alguna ventaja al jugador como por ejemplo puntos de victoria. Ahora, en la isla de Catán existen losetas, representando 5 materias primas. Estos materiales son arcilla, madera, paja, roca y borrego. Para poder construir una aldea necesitas pagar una madera, una arcilla, una paja y un borrego. Una carretera cuesta una madera y una arcilla, y una ciudad cuesta dos pajas y tres rocas. Finalmente una carta de desarrollo cuesta una paja, una roca y un borrego. Estas materias primas se consiguen al principio de cada turno de un jugador donde se tiran dos dados de 6 caras. Cada loseta (por ende materia prima) tiene un número asignado. Cuando la suma de los dados cae en dicho número, y el jugador tiene una aldea o ciudad en una encrucijada con el recurso, el jugador entonces recibe esa materia prima. Si es una aldea recibe 1 materia prima, y si es una ciudad recibe 2 de la materia prima. Por ejemplo si en los dados cayó un 6, la loseta de arcilla tiene un 6 y el jugador tiene una aldea en la encrucijada colindante con esa arcilla, ese jugador recibe una arcilla durante ese turno. El juego recomienda un cierto acomodo de las losetas y números (ya que se pueden colocar las losetas de forma aleatoria). Este acomodo es el siguiente:

Ahora algo interesante del juego es el acomodo inicial de las aldeas (también llamado la fase de desarrollo). Después de acomodar las losetas y las fichas con los números, el primer jugador tiene la opción de acomodar su aldea en cualquier lugar de la isla de Catán. Yo voy a analizar cuál es el lugar más conveniente para el jugador poner su primera aldea para que en torno pueda construir más aldeas, tomando como base el acomodo que recomienda el juego.

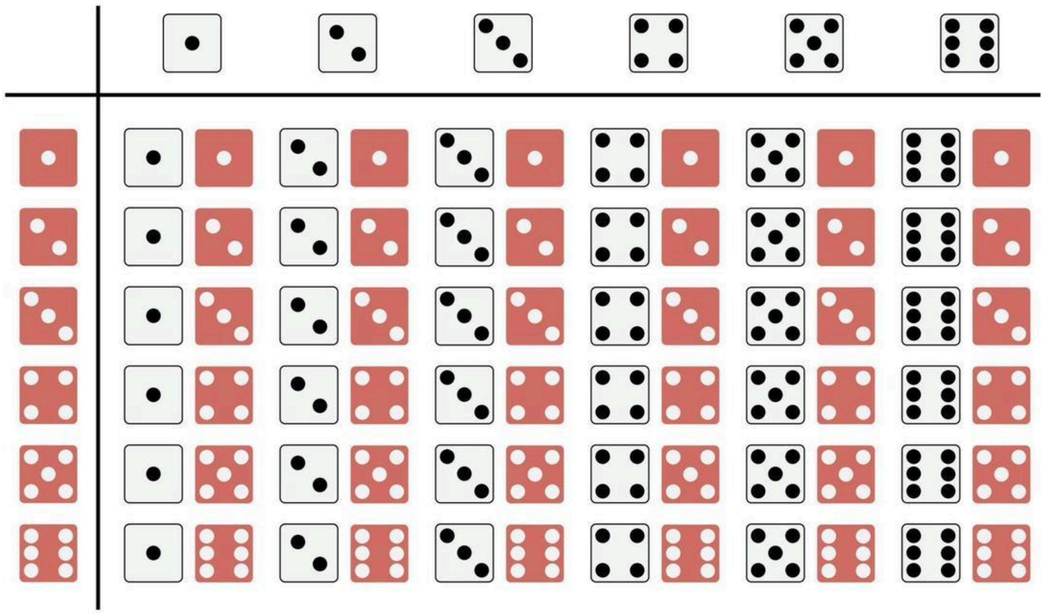
Descripción de las variables y su relación entre ellas (SEG0501 Pensamiento sistémico)

Variables Independientes:

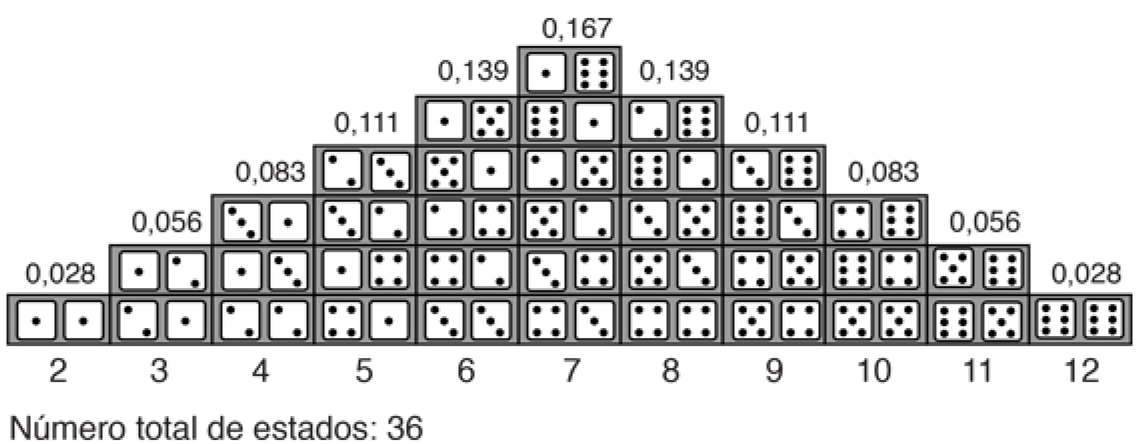
Las variables independientes que se van a utilizar son las siguientes:

1. Tirada de los dados:

   Los dados pueden conseguir un valor de 2 a 12, donde todas las combinaciones de los 2 dados de 6 siguen una distribución normal. En la siguiente imagen se muestran todas las tiradas posibles:



   De igual forma en el siguiente diagrama se muestra la distribución normal de los dados junto con la probabilidad que caiga cada número.



   Esta tabla es importante, ya que muestra qué tan probable es que salgan los números para poder conseguir materias primas.

2. Utilidad de encrucijada:

   Como había mencionado anteriormente, una aldea requiere madera, arcilla, paja y borrego para ser construida. Por lo mismo, la roca no es útil para el modelo que se va a realizar. De igual forma, tomando en cuenta que se necesita una diversidad de materias primas para construir, si en cierta encrucijada hay una materia prima repetida, la segunda de estas no se va a tomar en cuenta, ya que no es útil para avanzar lo más rápido posible. Por esta razón, se tiene **p(utilidad encrucijada)**, que se va a calcular con el número de materias primas útiles en la encrucijadas entre el número de materias primas adyacentes.

   a. **p(utilidad encrucijada)** = # materias útiles / materias adyacentes.
   b. Algunos ejemplos de esto es:
      i. **p(utilidad encrucijada)** = 2 (madera, arcilla y roca) / 3
      ii. **p(utilidad encrucijada)** = 3 (paja, madera y arcilla) / 3
      iii. **p(utilidad encrucijada)** = 2 (paja, paja y borrego) / 3
      iv. **p(utilidad encrucijada)** = 1 (madera, madera y roca) / 3

Variables Dependientes:

En este caso, la variable dependiente es predecir en qué encrucijada es más probable que un jugador consiga materias primas para que pueda expandirse y que pueda conseguir más aldeas en el futuro. Esta se va a definir como **p(éxito encrucijada).**

Por ejemplo, en la imagen se puede ver que se seleccionó la encrucijada con 9 para madera, 4 para borrego y 10 para arcilla. Lo que se quiere sacar es la probabilidad que el jugador en esa encrucijada tiene para conseguir las materias primas necesarias para que pueda construir aldeas.

Relación entre variables:

Una **encrucijada** tiene 3 **números asignados** donde cada número representa a una materia prima. Lo que se quiere entender es la probabilidad que cada encrucijada tiene para que un jugador pueda conseguir materias primas necesarias para hacer aldeas a través de los **dados**. Es importante aquí mencionar que se van a omitir las ciudades, carreteras y cartas de desarrollo en este análisis para hacerlo más simplificado. En otras palabras, queremos saber cuál es la mejor encrucijada (cuando todas están libres) para iniciar el juego, que tenga la probabilidad de conseguir más materias primas para construir aldeas.

Explicación del modelo (SING0202 Interpretación de variables)

(SEG0501 Pensamiento sistémico)

Ahora bien, antes de poder plantear qué encrucijada es mejor colocarte para poder construir aldeas, primero es necesario sacar la probabilidad de cada encrucijada individualmente, sin tomar en cuenta las materias primas con la que interactúa la encrucijada.

Las encrucijadas que voy a calcular su probabilidad se encuentran en la siguiente imagen, donde son un total de 18 encrucijadas.

Hay más que 18 encrucijadas en el tablero, pero todas las restantes están en la orilla (donde solo interactúan dos losetas) o interactúan con el desierto, que no tiene asignado un número (por lo que ahí también solo interactúan dos losetas).

Al tomar en cuenta que estos números representan una tirada de dados, en una encrucijada se toman cualquiera de los tres números. La cosa importante es que se toma en cuenta uno de los tres números, y no se necesita que caigan los tres. Por eso la probabilidad de cada número se suma en cada encrucijada. Como se suman, no es relevante tomar las encrucijadas que solo tengan dos números, ya que la probabilidad de dichas encrucijadas siempre va a ser menor que una de tres. Esta es la razón por la que decidí usar las 18 encrucijadas del centro, e ignorar para mi modelo todas las restantes.

Lo que ahora voy a hacer, es una tabla con las 18 encrucijadas, y se van a poder observar los 3 números que tiene cada encrucijada. Para eso, voy a usar la numeración que se encuentra en la foto.

	Número 1	Número 2	Número 3
Encrucijada 1	12	10	6
Encrucijada 2	10	2	6
Encrucijada 3	6	2	4
Encrucijada 4	2	4	9
Encrucijada 5	4	9	10
Encrucijada 6	4	10	3
Encrucijada 7	10	3	8
Encrucijada 8	3	8	5
Encrucijada 9	3	4	5
Encrucijada 10	4	5	11
Encrucijada 11	6	4	11
Encrucijada 12	3	4	6
Encrucijada 13	3	5	6
Encrucijada 14	8	3	5
Encrucijada 15	8	11	3
Encrucijada 16	8	9	11
Encrucijada 17	9	11	12
Encrucijada 18	12	6	11

Ahora bien, la tabla con los números no representa ninguna información útil. Como cada de esos números representa una tirada de dados, ahora lo que hace falta es sacar la probabilidad que representa cada tirada de dados. Primero en la siguiente tabla pongo todas las combinaciones de los dados, con su resultado, y hago un conteo de cuántas veces aparece dicho resultado.

Dado 1	Dado 2	Resultado	Conteo
1	1	2	1
1	2	3	2
2	1		
1	3	4	3
3	1		
2	2		
1	4	5	4
4	1		
2	3		
3	2		
5	1	6	5
1	5		
3	3		
2	4		
4	2		
1	6	7	6
6	1		
2	5		
5	2		
3	4		
4	3		
2	6	8	5
6	2		
4	4		
3	5		
5	3		
6	3	9	4
3	6		
4	5		
5	4		
4	6	10	3
6	4		
5	5		
5	6	11	2
6	5		
6	6	12	1

Después tomando como base el conteo que aparece cada resultado, se puede generar la siguiente tabla, donde se observa la probabilidad de cada tirada de dados (del 2 al 12)

Resultado de dado	Conteo	Probabilidad	Porcentaje
2	1	0.0278	2.78%
3	2	0.0556	5.56%
4	3	0.0833	8.33%
5	4	0.1111	11.11%
6	5	0.1389	13.89%
7	6	0.1667	16.67%
8	5	0.1389	13.89%
9	4	0.1111	11.11%
10	3	0.0833	8.33%
11	2	0.0556	5.56%
12	1	0.0278	2.78%

Finalmente, tomando la probabilidad calculada de cada número, se creó una nueva tabla sumando las 3 probabilidades de :

	Número 1	Prob #1	Número 2	Prob #2	Número 3	Prob #3	Probabilidad total	Porcentaje total
Encrucijada 1	12	0.0278	10	0.0833	6	0.1389	0.2500	25.00%
Encrucijada 2	10	0.0833	2	0.0278	6	0.1389	0.2500	25.00%
Encrucijada 3	6	0.1389	2	0.0278	4	0.0833	0.2500	25.00%
Encrucijada 4	2	0.0278	4	0.0833	9	0.1111	0.2222	22.22%
Encrucijada 5	4	0.0833	9	0.1111	10	0.0833	0.2778	27.78%
Encrucijada 6	4	0.0833	10	0.0833	3	0.0556	0.2222	22.22%
Encrucijada 7	10	0.0833	3	0.0556	8	0.1389	0.2778	27.78%
Encrucijada 8	3	0.0556	8	0.1389	5	0.1111	0.3056	**30.56%**
Encrucijada 9	3	0.0556	4	0.0833	5	0.1111	0.2500	25.00%
Encrucijada 10	4	0.0833	5	0.1111	11	0.0556	0.2500	25.00%
Encrucijada 11	6	0.1389	4	0.0833	11	0.0556	0.2778	27.78%
Encrucijada 12	3	0.0556	4	0.0833	6	0.1389	0.2778	27.78%
Encrucijada 13	3	0.0556	5	0.1111	6	0.1389	0.3056	**30.56%**
Encrucijada 14	8	0.1389	3	0.0556	5	0.1111	0.3056	**30.56%**

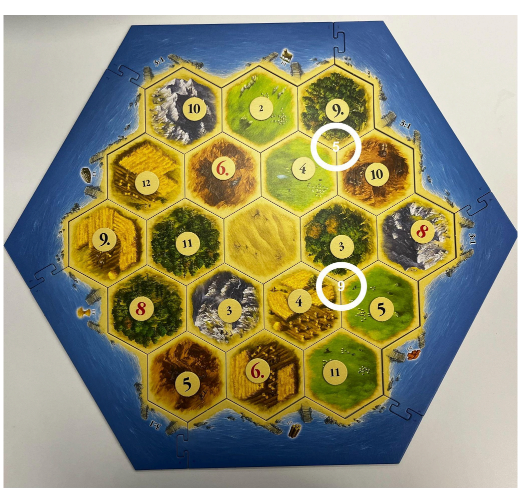
Encrucijada 15	8	0.1389	11	0.0556	3	0.0556	0.2500	25.00%
Encrucijada 16	8	0.1389	9	0.1111	11	0.0556	0.3056	**30.56%**
Encrucijada 17	9	0.1111	11	0.0556	12	0.0278	0.1944	19.44%
Encrucijada 18	12	0.0278	6	0.1389	11	0.0556	0.2222	22.22%

Los resultados fueron que las encrucijadas 8, 13, 14 y 16 son las que tienen la probabilidad más alta de salir con un 30.56% de probabilidad de sacar uno de esos números.

Vuelvo a anexar la imagen para poder visualizar las encrucijadas y sus materias primas.

Ahora lo importante a notar es que esto solo representa a las encrucijadas sin tomar en cuenta la materia prima que estás recibiendo. Considerando que lo que quiero modelar, es la mejor encrucijada para poder construir aldeas, tengo que incluir a **p(utilidad encrucijada)** en cada una de las encrucijadas. Por ejemplo, la encrucijada 1 tendría una utilidad de ⅔, ya que tiene roca, paja y arcilla, y solo son útiles la paja y arcilla. De igual forma si una encrucijada tiene una materia prima repetida, la materia repetida no se toma en cuenta como útil, ya que se necesita diversidad para poder avanzar. Esta probabilidad se saca para todas las encrucijadas, y se multiplica por la probabilidad que la encrucijada tiene de salir en los dados (ya que en este caso queremos que ambas se cumplan). Esto se planteó en la siguiente tabla:

	p(encrucijada)	Materiales encrucijada	p(utilidad encrucijada)	p(exito encrucijada)	Porcentaje
Encrucijada 1	0.2500	Roca, paja y arcilla	0.67	0.1666666667	16.67%
Encrucijada 2	0.2500	Roca, arcilla y borrego	0.67	0.1666666667	16.67%
Encrucijada 3	0.2500	Arcilla, borrego y borrego	0.67	0.1666666667	16.67%
Encrucijada 4	0.2222	Borrego, borrego y madera	0.67	0.1481481481	14.81%
Encrucijada 5	0.2778	Madera, borrego y arcilla	1.00	0.2777777778	**27.78%**
Encrucijada 6	0.2222	Borrego, madera y arcilla	1.00	0.2222222222	22.22%
Encrucijada 7	0.2778	Arcilla, madera y roca	0.67	0.1851851852	18.52%
Encrucijada 8	0.3056	Madera, roca y borrego	0.67	0.2037037037	20.37%
Encrucijada 9	0.2500	Madera, paja y borrego	1.00	0.25	**25.00%**
Encrucijada 10	0.2500	Paja, borrego y borrego	0.67	0.1666666667	16.67%
Encrucijada 11	0.2778	Paja, paja y borrego	0.67	0.1851851852	18.52%
Encrucijada 12	0.2778	Roca, paja y paja	0.33	0.09259259259	9.26%
Encrucijada 13	0.3056	Roca, arcilla y paja	0.67	0.2037037037	20.37%
Encrucijada 14	0.3056	Madera, roca y arcilla	0.67	0.2037037037	20.37%
Encrucijada 15	0.2500	Madera, madera y roca	0.33	0.08333333333	8.33%
Encrucijada 16	0.3056	Paja, madera y madera	0.67	0.2037037037	20.37%
Encrucijada 17	0.1944	Paja, paja y madera	0.67	0.1296296296	12.96%
Encrucijada 18	0.2222	Paja, arcilla y madera	1.00	0.2222222222	22.22%



Se puede determinar entonces que **p(éxito encrucijada)** tomando como base la disposición sugerida del juego, es más alta en la encrucijada 5, seguida con la encrucijada 9.

El modelo para llegar a ese resultado consiste en:

p(éxito encrucijada) = p(encrucijada) * p(utilidad encrucijada)

Donde:

p(encrucijada) = p(numero 1) + p(numero 2) + p(numero 3)

p(numero) = p(dado)

p(utilidad encrucijada) = número de materias útiles / número de materias adyacentes

Estrategia sugerida siguiendo el modelo (SING0202 Interpretación de variables)
Siguiendo el modelo entonces se puede determinar que la encrucijada que más le conviene a un jugador es la encrucijada 5, que tiene como materias primas a madera (9), arcilla (10) y borregos (4), seguida por la encrucijada 9 que tiene como materias primas a madera (3), paja (4) y borregos (5).

Demostración de uso del modelo (SING0202 Interpretación de variables)

En este video https://youtu.be/53viAwSELFs se puede observar como se usa el modelo, todo basado en la siguiente tabla que se sacó anteriormente:

	p(encrucijada)	Materiales encrucijada	p(utilidad encrucijada)	p(exito encrucijada)	Porcentaje
Encrucijada 5	0.2778	Madera, borrego y arcilla	1.00	0.2777777778	**27.78%**
Encrucijada 9	0.2500	Madera, paja y borrego	1.00	0.25	**25.00%**