Ejercicios ejemplo de probabilidad

Rodrigo Maranzana

Falacia del apostador

Chicago Booth, Gamble's Fallacy: *https://www.youtube.com/watch?v=4eVluL-idkM*

Falacia del apostador

Ejercicio 00: Distribución Normal

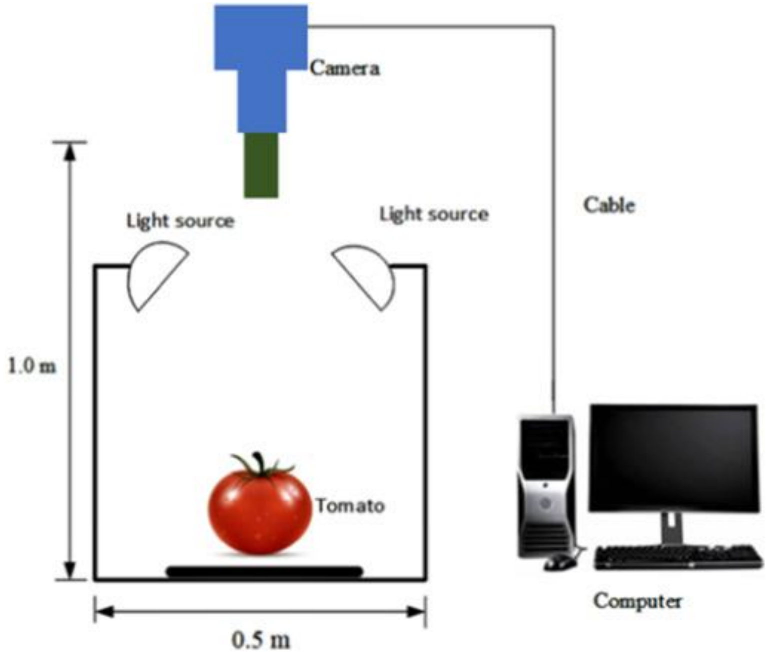
Una empresa comercializa tomates cherry premium. En su línea de producción automatizada, una unidad con un pistón y una cámara conectada a un modelo de visión artificial, selecciona tomates de entre 3.3 y 3.9 cm de diámetro; valor que la compañía considera aceptable para ser considerados de esa gama.

Se sabe que la media de los tomates es de 3.5cm y el desvío 0.35 cm. Además, se sabe que el diámetro en cada batch es una variable aleatoria que sigue una distribución normal.

Una nueva oportunidad de negocio, obliga a ingeniería a dimensionar un depósito de productos seleccionados bajo las anteriores características.

Sabemos que el próximo batch de producción incluye 153000 unidades variadas, se busca saber qué cantidad de esas unidades se almacenarán como premium. En base a ese dato, se alquilará un nuevo depósito.

$$\bar{X} = 3.5cm,$$
$$\sigma = 0.35cm,$$
$$X_{min} = 3.3cm,$$
$$X_{max} = 3.9cm,$$
$$q = 153000$$



Ireri et. Al (2019) – "A computer vision system for defect discrimination and grading in tomatoes using machine learning and image processing"

Ejercicio 00: Distribución Normal

$$Calculamos\ los\ Z - Scores:$$

$$Z = \frac{X - \bar{X}}{\sigma}$$

$$Zmin = \frac{3.3 - 3.5}{0.35} = -0.5714$$

$$Zmax = \frac{3.9 - 3.5}{0.35} = 1.1428$$

$$P(3.3cm \leq X \leq 3.9cm)$$

$$P(-0.5714 \leq X \leq 1.1428)$$

$$P(X \leq 1.1428) - P(X \leq -0.5714)$$

$$P(X \leq 1.1428) = 0.87345$$

$$P(X \leq -0.5714) = 0.28385$$

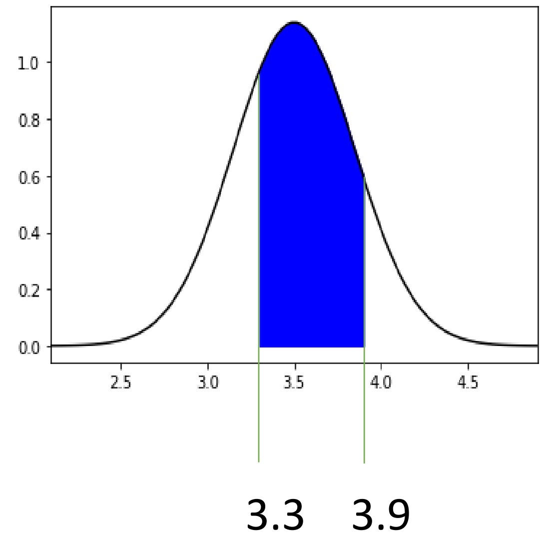
$$P(X \leq 1.1428) - P(X \leq -0.5714) = \mathbf{0.58959}$$

$$P(3.3cm \leq X \leq 3.9cm) * q$$

$$= 0.58959\ * 153000$$

$$= \mathbf{90209}$$

Cantidad a stockear para dimensionamiento.



Ejercicio 01: Distribución Poisson

Siguiendo con el caso anterior. En mantenimiento industrial, surge la necesidad de presupuestar mensualmente los servicios de reparación correctivos del robot seleccionador.
Particularmente nos interesa centrarnos en el pistón. Sabemos que el robot tiene una media de fallas graves de 1 cada 20 días por desajuste del pistón. Se trabaja 24 días al mes. La cantidad de fallas es una variable que sigue una distribución Poisson.
Al ocurrir por lo menos dos fallas, el servicio de mantenimiento para la línea y hace los ajustes correspondientes.
El costo de reparación es de 500 dólares.

$$\mu = 1\,\frac{falla}{20\,días} * 24\,\frac{días}{mes} = 1.2\ {}^{fallas}/_{mes}$$

$$costo = 500\,usd$$

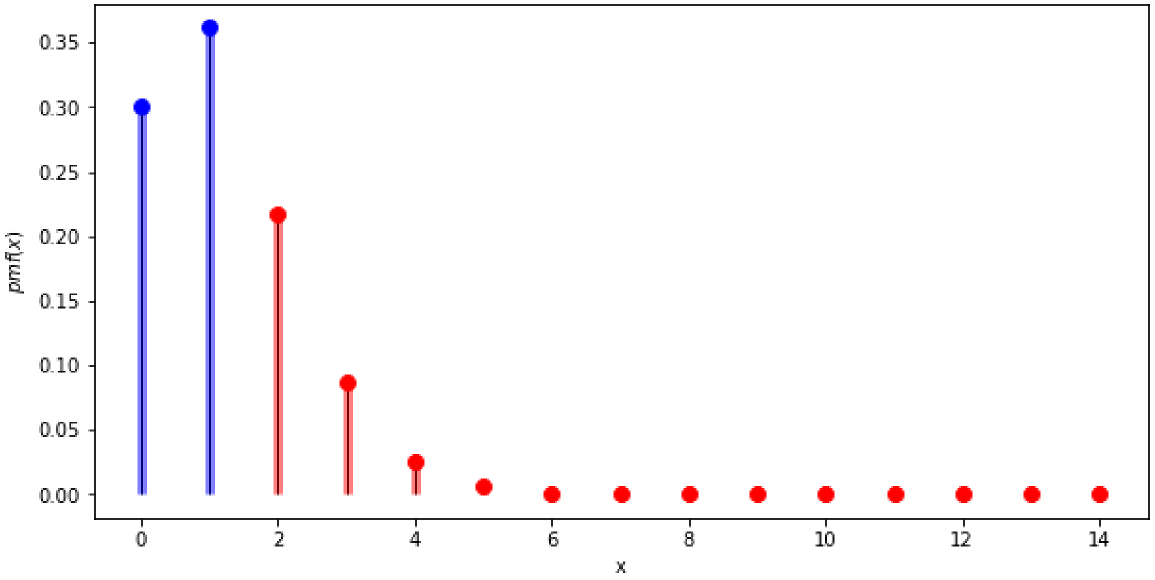
Ejercicio 01: Distribución Poisson

$$P(X \geq 2) = \quad 1\ - [P(X = 0) + P(X = 1)]$$

$$P(X \geq 2) = 1 - (0.30119 + 0.36143)$$

$$\boldsymbol{P(X \geq 2) = 0.33737}$$

$$P(x) = \frac{e^{-\lambda}\lambda^{x}}{x!}$$



Presupuesto: $P(X \geq 2) * \text{costo} = 0.33737\ * 500\text{usd}$
$= \mathbf{168.69\ usd}$

Ejercicio 02: Repaso de probabilidad

Una empresa fabrica y comercializa productos impresos en 3D por su web enlazada a un sitio e-commerce.

Con el objetivo de evaluar el funnel de viaje del cliente a través de la página web, se obtienen los siguientes datos en una ventana de tiempo:

- Cantidad de usuarios en la landing page: 8590
- Porcentaje de usuarios que acceden a publicación sample: 11,58%
- Procentaje de usuarios que cliquean “comprar” en la publicación sample: 36,18%
- Porcentaje de usuarios que terminan el proceso: habiendo cliqueado comprar: 9,72%

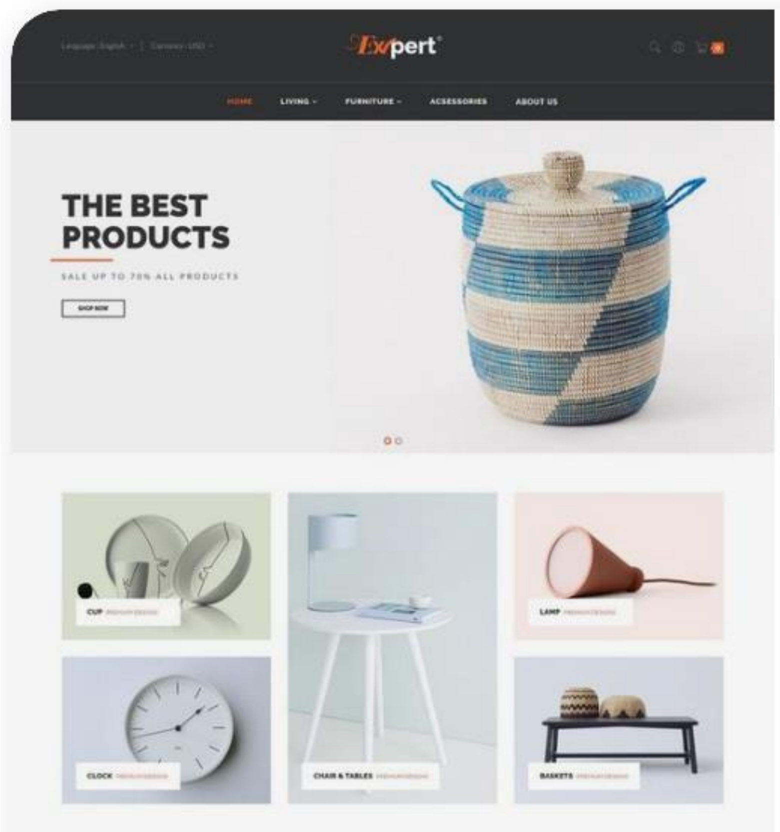


Imagen ejemplo, themeforest.net

Ejercicio 02: Repaso de probabilidad

Landing Page: **8590 usuarios**

Funnel

Publicación sample: **11,58%**

Proceso de compra: **36,18%**

Compra exitosa: **9,72%**

1. Probabilidad de entrar a publicación sample: $P(publicación) = 11{,}58\%$

2. Probabilidad de iniciar proceso de compra tal que ingresó a sample:

$$P(compra|publicación) = 36{,}18\%$$

3. Probabilidad de compra exitosa tal que compró: $P(exitosa|compra) = 9{,}72\%$

4. Probabilidad de deserción en proceso de compra:

$$P(deserción|compra) = 1 - P(exitosa|compra) = 90{,}28\%$$

Ejercicio 02: Repaso de probabilidad

Landing Page: **8590 usuarios**

Funnel

Publicación sample: **11,58%**

Proceso de compra: **36,18%**

Compra exitosa: **9,72%**

5. Probabilidad de compra exitosa tal que se entró en la publicación:

$$P(exitosa|publicación) =$$
$$P(exitosa \cap compra|publicación) =$$
$$P(exitosa|compra) * P(compra|publicación) =$$
$$0{,}0972 * 0{,}3618 = 0{,}0335 = 3{,}35\%$$

6. Cantidad de usuarios de compra exitosa:

$$P(exitosa \cap publicación) =$$
$$P(exitosa|publicación) * P(publicación) =$$
$$0{,}1158 * 0{,}0335 = 0{,}00407 = \mathbf{0{,}407}\%$$

$$P(exitosa) * Q_{landing} = 0{,}407\% * 8590 \text{usuarios} = 35 \text{ usuarios}$$

Ejercicio 03: Distribución Gamma

Siguiendo el ejemplo anterior:

Si bien la página tiene la clásica estructura de catálogo de productos en mosaico; los diseñadores optaron por mostrar renders 3D embebidos junto a la información y precio de cada producto. Esta decisión tiene que ver con que algunos productos no están pintados y se dificulta su visualización en una imagen estática.

Sin embargo, depende del dispositivo y la conexión de cada cliente que la experiencia de usuario resulte realmente exitosa.

Se decidió analizar el (click-through rate) CTR a través del pipeline de publicaciones. Llegando a la conclusión que los clientes con mala conexión desertaban al intentar visualizar un producto; claramente impulsado por la carga del render.

Ejercicio 03: Distribución Beta

Por lo tanto, se decidió:

- Evitar cargar la renderización para los usuarios con latencias por encima de los 500ms; reemplazándolo por una imagen estática.
- Cargar un .gif para usuarios con latencias entre 350 y 500ms.
- Por debajo de 350ms seguir cargando la renderización.

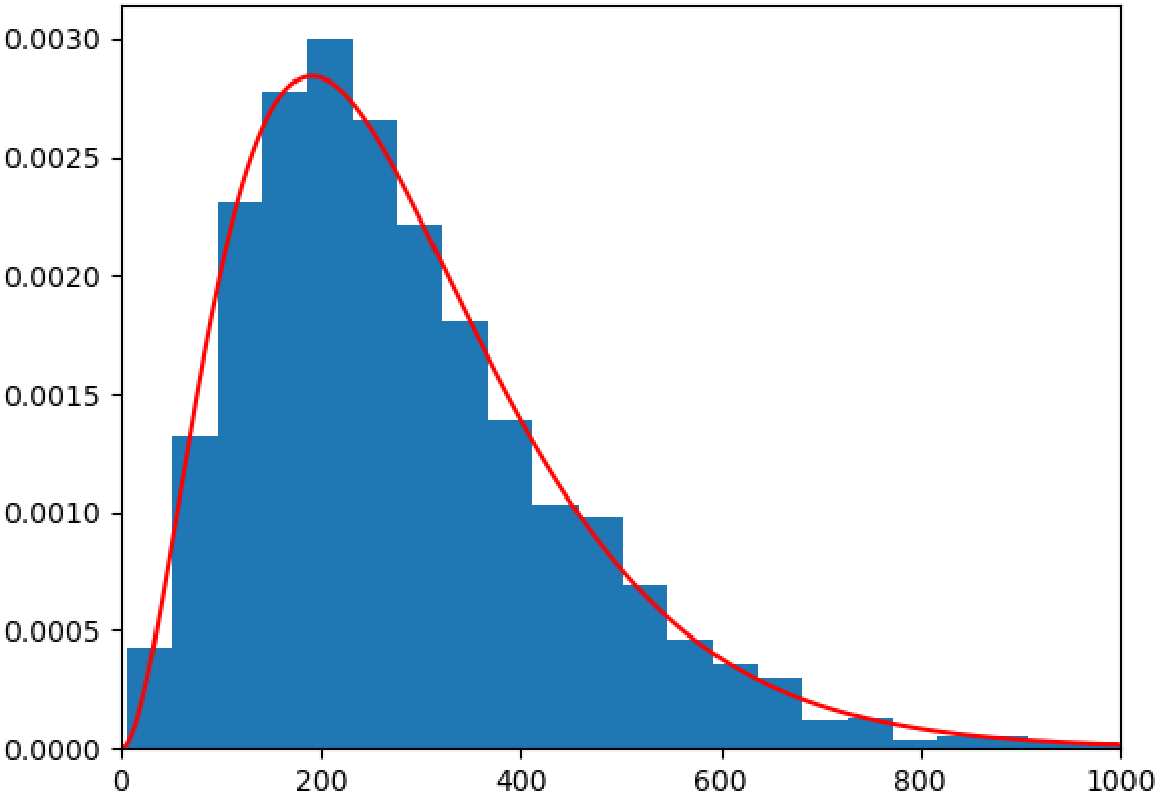
Se sabe que históricamente los usuarios se tienen una media de 200ms y se supone que siguen una distribución de Gamma. Se espera que el tráfico diario sea de 2000 usuarios.

Con el objetivo de dimensionar el sistema antes de presupuestarlo; se pide calcular la cantidad diaria de usuarios que estarían en cada grupo.

Ejercicio 03: Distribución Gamma

Se hizo un estudio de las latencias y se ajustó la distribución Gamma a los valores históricos, obteniendo el siguiente resultado:

<u>Parámetros:</u>
K = 3
λ = 0.0105



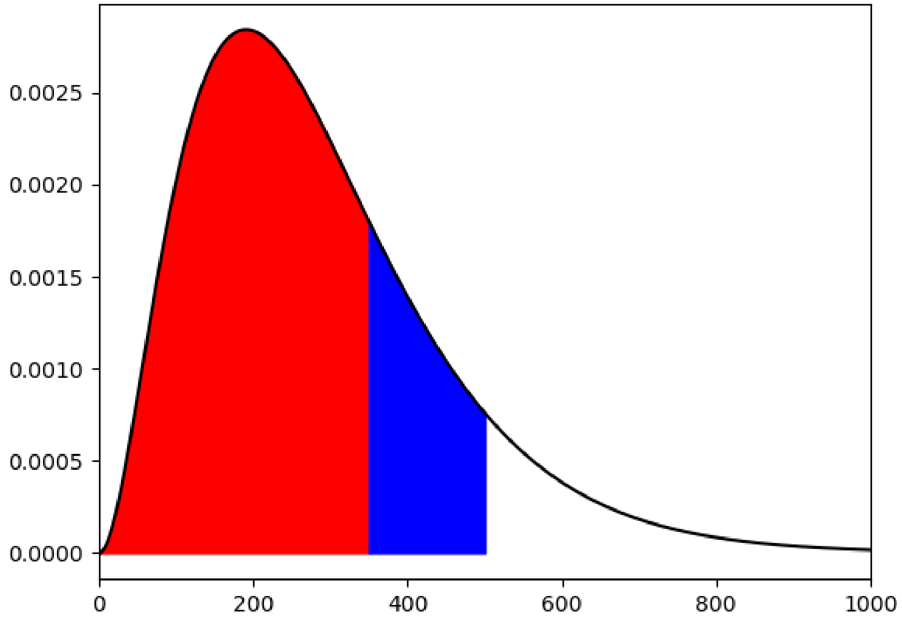
Ejercicio 03: Distribución Gamma

$P(X \leq 350ms)$

$P(350ms \leq X \leq 500ms) \;=\; P(X \leq 500ms) - \mathrm{P}(\mathrm{X} \leq 350\mathrm{ms})$

$P(X \geq 500ms) \;= 1 - P(X \leq 500ms)$

Si $X \sim \Gamma(k = 3, \lambda = 0.0105)$



Render: $P(X \leq 350ms) * Q_{Tot} = 0.7103 * 2000 = 1421$

Gif: $P(350ms \leq X \leq 500ms) * Q_{Tot} = 0.1845 * 2000 = 370$

Imagen: $P(X \geq 500ms) * Q_{Tot} = 0.1051 * 2000 = 211$