VE401 Recitation 2

Discrete Random Variable

Discrete Random Variable and Probability Density Function (PDF)

A ***discrete random variable*** X maps the sample space to a countable subset Ω of $\mathbb{R}$, with each number representing an event, and the ***probability density function*** f_X maps the subset to its probability. The PDF must follow

- $f_X(x) > 0$ for all x.
- $\sum_{x\in\Omega} f_X(x) = 1$.

Note: For various distribution of random variable X, we need to know its

- parameter(s),
- $\mathrm{E}[X]$,
- $\mathrm{Var}[X]$,
- probability density function (PDF),
- cumulative distribution function (CDF),
- moment generating function (MGF), and
- when to use it?

Expectation

Intuition: given a set of data, what would I expect the next number to be?

Mathematical representation: for discrete random variable, $\mathrm{E}[X] = \mu_X = \mu = \sum_{x\in\Omega} x\, f_X(x)$.

Properties: for any random variable X and Y,

- For a constant $c \in \mathbb{R}$, $\mathrm{E}[c] = c$, $\mathrm{E}[c\,X] = c\,\mathrm{E}[X]$,
- $\mathrm{E}[X+Y] = \mathrm{E}[X] + \mathrm{E}[Y]$,
- For any function $\varphi : \Omega \to \mathbb{R}$, $\mathrm{E}[\varphi\circ X] = \sum_{x\in\Omega} \varphi(x)\, f_X(x)$.

What do these properties imply?

$\mathrm{E}[\cdot]$ is a linear operation!

Variance and Standard Variance

Intuition: how much does random variable deviate from the mean?

Mathematical representation:

- Variance: $\mathrm{Var}[X] = \sigma_X{}^2 = \sigma^2 = \mathrm{E}\left[\left(X - \mathrm{E}[X]\right)\right]^2 = \mathrm{E}\left[X^2\right] - \mathrm{E}[X]^2$.
- Standard variance: $\sigma_X = \sqrt{\mathrm{Var}[X]}$.

Properties:

- For a constant $c \in \mathbb{R}$, $\mathrm{Var}[c] = 0$, $\mathrm{Var}[c\,X] = c^2\,\mathrm{Var}[X]$.
- For independent random variable X and Y, $\mathrm{Var}[X + Y] = \mathrm{Var}[X] + \mathrm{Var}[Y]$.

Standardization

We can standardize a random variable X by calculating $Y = \frac{X-\mu}{\sigma}$. It has mean $\mu_Y = 0$ and standard deviation $\sigma_Y = 1$.

Example:

Show that $\sigma_Y = 1$.

We have $\mathrm{Var}[Y] = \frac{1}{\sigma^2}\mathrm{Var}[X - \mu] = \frac{1}{\sigma^2}(\mathrm{Var}[X] + 0) = \frac{1}{\sigma^2}\mathrm{Var}[X] = 1$.

Moment and Moment Generating Function (MGF)

Intuition: MGF encodes the sequence of all ***moments***, $\mathrm{E}\left[X^0\right]$, $\mathrm{E}\left[X^1\right]$, $\mathrm{E}\left[X^2\right]$, ... into one function.

Mathematical representation: MGF exists iff $\mathrm{E}\left[e^{t\,X}\right]$ exists, in which case

$$m_X(t) = \sum_{k=0}^{\infty} \frac{\mathrm{E}\left[X^k\right]}{k!}\,t^k = \mathrm{E}\left[e^{t\,X}\right]$$

and the k^{th} moment can be calculated using $\mathrm{E}\left[X^k\right] = \left(\frac{d^k\,m_X(t)}{d\,t^k}\right)_{t=0}$.

Properties: Assume MGF exists,

- if two distributions have same MGF, then two distributions are identical. (see assignment)
- For any constant α, $\beta \in \mathbb{R}$, $m_{\alpha\,X+\beta}(t) = e^{\beta\,t}\,m_X(\alpha\,t)$.
- For independent random variable X_1, X_2, $m_{X_1+X_2} = m_{X_1}\,m_{X_2}$.

Example:

Discrete Random Variable X has moment-generating function $m_X(t) = \frac{1}{6}\,e^{-2\,t} + \frac{1}{3}\,e^{-t} + \frac{1}{4}\,e^{t} + \frac{1}{4}\,e^{2\,t}$, find $P[\,|X| \le 1]$.

The general MGF of a discrete RV is $m_X(t) = \sum_{x\in\Omega} f_X(x)\, e^{t\,x}$. From this we can calculate the actual PDF is

$$f_X(x) = \begin{cases} \frac{1}{6} & \text{if } x = -2 \\ \frac{1}{3} & \text{if } x = -1 \\ \frac{1}{4} & \text{if } x = 1 \text{ or } 2 \\ 0 & \text{otherwise} \end{cases}$$

Thus $P[\,|X|\, \leq 1] = P[X = -1] + P[X = 1] = \frac{1}{3} + \frac{1}{4} = \frac{7}{12}$.

Given that the Normal Distribution (μ,σ) has MGF $m_X(t) = e^{\frac{\sigma^2 t^2}{2} + \mu\, t}$, use the properties of MGF to prove that $Y = \frac{X-\mu}{\sigma}$ also follows normal distribution.

We can easily calculate $Y = \frac{X-\mu}{\sigma} = \frac{1}{\sigma} X - \frac{\mu}{\sigma}$, we have $m_Y(t) = e^{-\frac{\mu}{\sigma} t}\, m_X\!\left(\frac{1}{\sigma}\, t\right) = e^{-\frac{\mu}{\sigma} t}\, e^{\frac{t^2}{2} + \mu \frac{t}{\sigma}} = e^{\frac{t^2}{2}}$.

We also know that normal random variable with $\mu = 0$ and $\sigma = 1$ has MGF $m_Z(t) = e^{\frac{t^2}{2}}$. So Y and Z has the same distribution (both standard normal).

Cumulative Distribution Function (CDF)

Bernoulli Distribution

Binomial Distribution

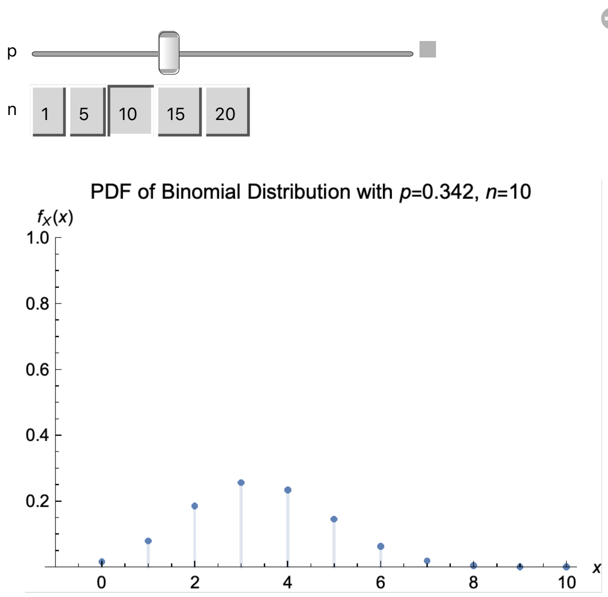
Purpose: how many successes in n trial(s)?

Parameter and properties:

- $p \in [0,\ 1]$, describing the probability of success in each trial. $q := 1 - p$.
- $n \in \{0,\ 1,\ 2,\ \ldots\}$ is the number of trials.

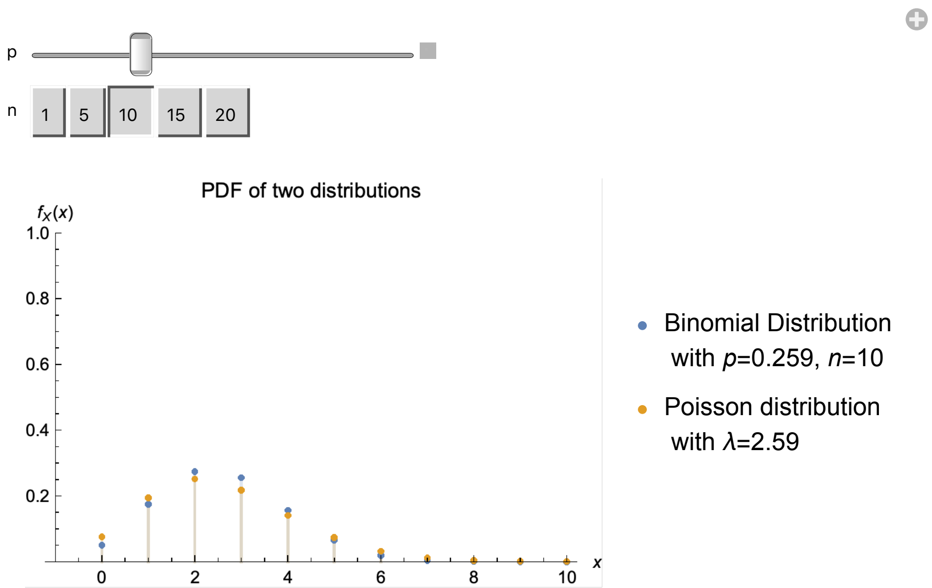
E[X]	Var[X]	PDF	MGF
$n\,p$	$n\,p\,q$	$\begin{cases} \binom{n}{x} p^x q^{n-x} & 0 \leq x \leq n \\ 0 & \text{otherwise} \end{cases}$	$\left(p\left(e^t - 1\right) + 1\right)^n$

Out[148]=



Approximation of PDF: if n is large and p is small, the PDF can be approximated by Poisson distribution with parameter $\lambda = n\,p$.

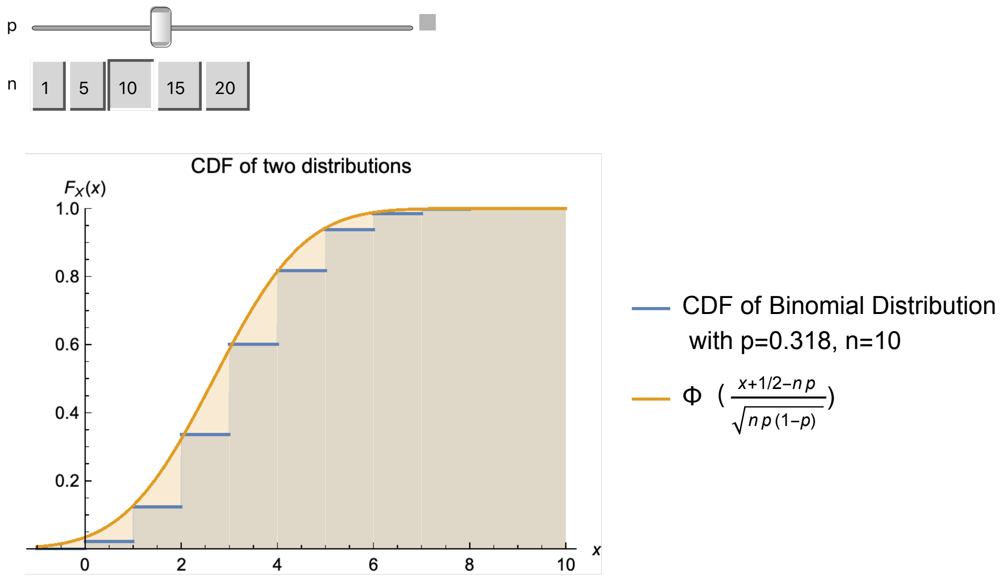
Out[ ]=



Approximation of CDF: if $\begin{cases} n > 10 & \text{if } p \text{ is close to } 1/2 \\ n\,p > 5 & \text{if } p \leq 1/2 \\ n(1-p) > 5 & \text{if } p > 1/2 \end{cases}$, the CDF at $y \in \mathbb{N}$ can be approximated by normal distribution.

$$P[X \leq y] = \sum_{x=0}^{y} \binom{n}{x} p^x (1-p)^{n-x} \approx \Phi\left(\frac{y+1/2-n\,p}{\sqrt{n\,p(1-p)}}\right)$$

Out[80]=



Geometric Distribution

Pascal Distribution

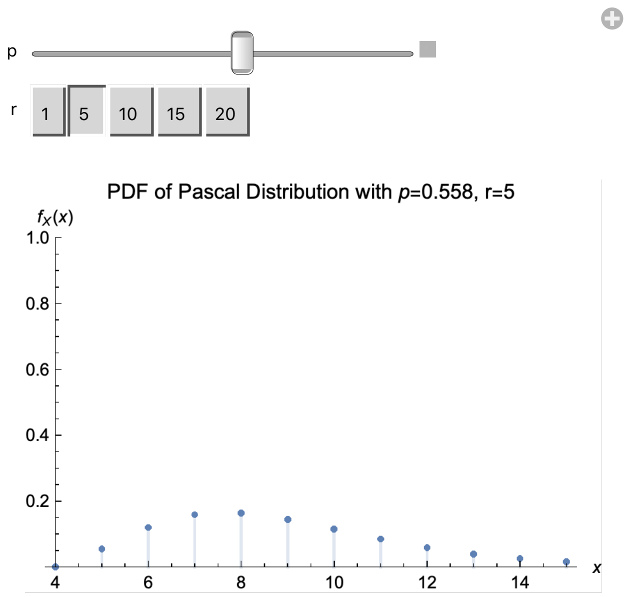
Purpose: the probability of having r^{th} success on x^{th} trial? $(x \geq r)$

Parameter and properties:

- $p \in (0, 1]$, describing the probability of success in each trial. $q := 1 - p$.
- $r \in \{1, 2, \ldots\}$ is the index of success we need.

$\mathrm{E}[X]$	$\mathrm{Var}[X]$	PDF	MGF
$\frac{r}{p}$	$\frac{q\,r}{p^2}$	$\begin{cases} \binom{x-1}{r-1} p^r q^{x-r} & x \geq r \\ 0 & \text{otherwise} \end{cases}$	$\left(\frac{p\,e^t}{1-q\,e^t}\right)^r$

Out[154]=



Use this to solve Banach Matchbox Problem. Suppose a mathematician carries two matchboxes at all times: one in his left pocket and one in his right. Each time he needs a match, he is equally likely to take it from either pocket. Suppose he reaches into his pocket and discovers for the first time that the box picked is empty. If it is assumed that each of the matchboxes originally contained n matches, what is the probability that there are exactly k matches in the other box?
What parameters should you use?

We can set *success* as picking a match from the matchbox that will eventually become empty, and *failure* as picking a match from the other box.

The question then transform to the following: **what's the probability of having $(n+1)^{\text{th}}$ success on $(2n-k+1)^{\text{th}}$ trial** ($+1$ because we drew the empty box, meaning that we performed $n+1$ draws on the box)**?** We can then use the Pascal distribution with $r = n+1$ and $p = 1/2$:

$$P[X = 2n-k+1] = \binom{2n-k+1-1}{n+1-1}\left(\frac{1}{2}\right)^{n-k}\left(\frac{1}{2}\right)^{n} = \binom{2n-k}{n}\left(\frac{1}{2}\right)^{2n-k}$$

Negative Binomial Distribution

Purpose: number of failures before r successes?

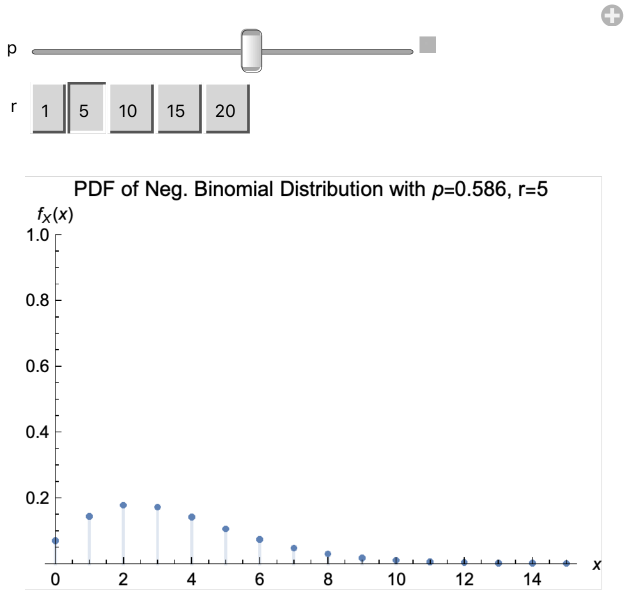
Parameter and properties:

- $p \in (0, 1]$, describing the probability of success in each trial. $q := 1-p$.
- $r \in \{1, 2, \ldots\}$ is the number of successes.

E[X]	Var[X]	PDF	MGF

$\frac{q\,r}{p}$	$\frac{q\,r}{p^2}$	$\begin{cases} \binom{r+x-1}{r-1} p^r (1-p)^x & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$	$\left(\frac{p}{1-q\,e^t}\right)^r$

Out[153]=



Example:

Use this to solve Banach Matchbox Problem. What parameters should you use?

We can set *success* as picking a match from the matchbox that will eventually become empty, and *failure* as picking a match from the other box.

The question then transform to the following: **what's the probability of having $n-k$ failures before $n+1$** ($+1$ because we drew the empty box, meaning that we performed $n+1$ draws on the box) **successes?** We can then use the negative binomial distribution with $r = n+1$ and $p = 1/2$:

$$P[X = n-k] = \binom{n-k+n+1-1}{n+1-1}\left(\frac{1}{2}\right)^{n-k}\left(\frac{1}{2}\right)^{n} = \binom{2\,n-k}{n}\left(\frac{1}{2}\right)^{2\,n-k}$$

Poisson Distribution

Purpose: how many arrivals in a small time interval?

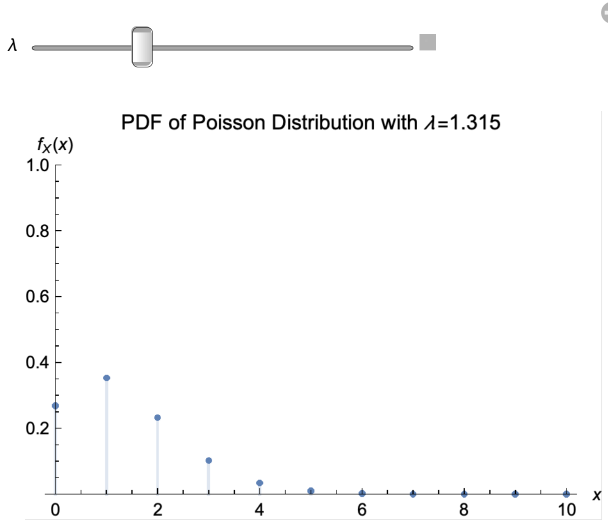
Parameter and properties:

- $\lambda > 0$, describing the rate of arrival.

Mean	Var[X]	PDF	MGF
λ	λ	$\begin{cases} \frac{e^{-\lambda}\lambda^x}{x!} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$	$e^{\lambda\,(e^t-1)}$

λ

Out[•]=



Example:

Raindrops keep falling on my head at an average rate of 20 drops/min. What is the probability of having no raindrop falling on my head in a given 3 seconds time interval?

The average rate $\lambda = 1$ (drop/3 seconds). $P[X = 0] = \frac{e^{-1} 1^0}{0!} = \frac{1}{e} = 36.7\,\%$.

Continuous Random Variables

Continuous Random Variable and its Properties

A ***continuous random variable*** is a map $X : S \to \mathbb{R}$ with a ***probability density function*** $f_X : \mathbb{R} \to \mathbb{R}$ such that

- $f_X(x) \geq 0$ for all $x \in \mathbb{R}$,
- $\int_{-\infty}^{\infty} f_X(x)\, dx = 1$.

And the ***cumulative distribution function*** is defined as

$$F_X(x) := P[X \leq x] = \int_{-\infty}^{x} f_X(y)\, dy\,,$$

which implies that $f_X(x) = F'_X(x)$.

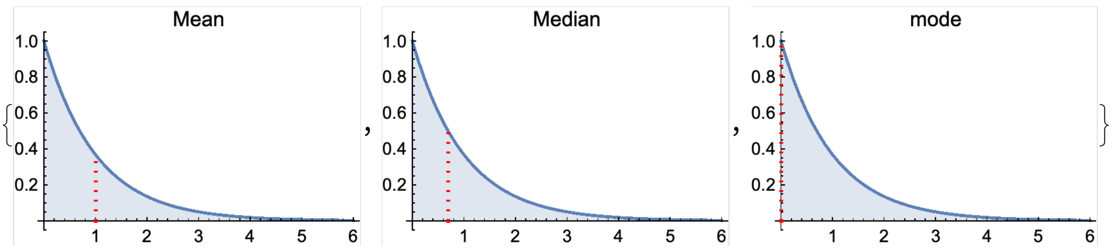
The ***expectation*** has the similar properties,

$$\mathrm{E}[X] := \int_{\mathbb{R}} x\, f_X(x)\, dx,\ \mathrm{E}[\varphi \circ X] = \int_{\mathbb{R}} \varphi(x)\, f_X(x)\, dx.$$

And the ***variance*** is given similarly, $\mathrm{Var}[X] = \mathrm{E}\left[\left(X - \mathrm{E}[X]\right)\right]^2 = \mathrm{E}\left[X^2\right] - \mathrm{E}[X]^2$.

The ***median*** M_X is a number such that $P[X \leq M_X] = 0.5$.

The ***mode*** is a number at which f_X is maximum.

Out[79]=


Exponential Distribution

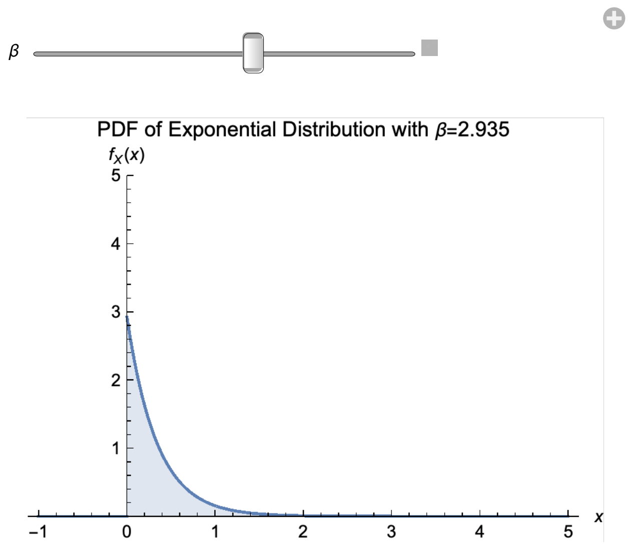
Purpose: time until first arrival?

Parameter and properties:

- $\beta > 0$ is the rate of arrival.

Mean	Variance	PDF	CDF	MGF
$\frac{1}{\beta}$	$\frac{1}{\beta^2}$	$\begin{cases} \beta\, e^{-\beta x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$	$\begin{cases} 1 - e^{-\beta x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$	$\frac{\beta}{\beta - t}$

- The exponential distribution is ***memoryless***, $P[X > x + s \mid X > x] = P[X > s]$.

Out[∘]=


Example:

Raindrops keep falling on my head at an average rate of 20 drops/min. What is the probability of having no raindrop falling on my head in a given 3 seconds time interval?

We can use $\beta = 1$ (drop/3 seconds). The probability of no drop (success) in the first 3 seconds is

$$P[X > 1] = 1 - \int_0^1 e^{-x}\, dx = 1 - \left(1 - \frac{1}{e}\right) = \frac{1}{e} = 36.7\,\%$$

Setting other values of β, such as $\beta = 20$ (drops/min) and calculating $P\left[X > \frac{1}{20}\right]$, will give the same result.

Given that no raindrop has fallen on my head in the past minute, what is the probability of having no raindrop on my head in another 3 seconds?

Still 36.7 % due to the memoryless property.

Gamma Distribution

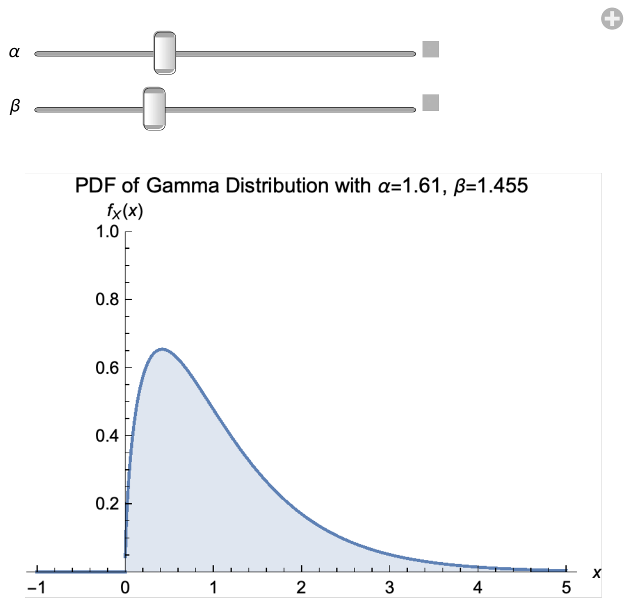
Purpose: time until α^{th} arrival?

Parameter and properties:

- $\alpha > 0$ is the number of arrivals you want.
- $\beta > 0$ is the rate of arrival.

Mean	Variance	PDF	MGF
$\frac{\alpha}{\beta}$	$\frac{\alpha}{\beta^2}$	$\begin{cases} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} & x > 0 \\ 0 & \text{otherwise} \end{cases}$ where $\Gamma(\alpha)=\int_0^\infty z^{\alpha-1} e^{-z}\, dz$	$\left(1-\frac{t}{\beta}\right)^{-\alpha}$

Out[138]=



Chi-squared Distribution

Purpose: how is the sum of squares of γ independent standard normal random variables distributed?

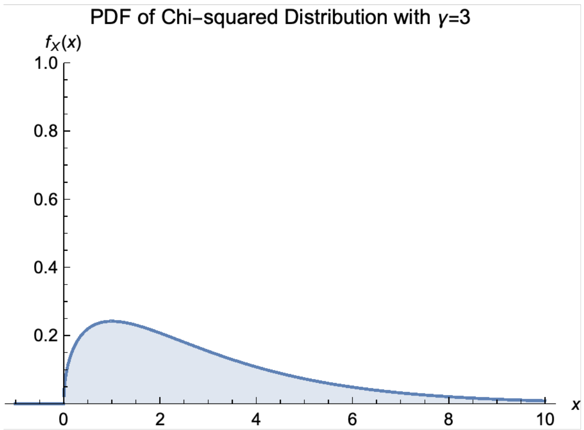
Parameter and properties:

- $\gamma \in \mathbb{N}^+$ is the ***degrees of freedom***.

Mean	Variance	PDF	MGF
γ	γ^2	$\begin{cases} \frac{1}{\Gamma\left(\frac{\gamma}{2}\right) 2^{\gamma/2}} x^{\frac{\gamma}{2}-1} e^{-x/2} & x > 0 \\ 0 & \text{True} \end{cases}$	$(1-2t)^{-\gamma/2}$

γ

Out[•]=



Transformation of Random Variable

X is a continuous RV and $\varphi : \mathbb{R} \to \mathbb{R}$ is a strictly monotonic and differentiable function, then the PDF of $Y = \varphi \circ X$ is

$$f_Y(y) = \begin{cases} f_X\left(\varphi^{-1}(y)\right) \cdot \left| \frac{d\varphi^{-1}(y)}{dy} \right| & \text{for } y \in \operatorname{ran} \varphi \\ 0 & \text{otherwise} \end{cases}$$

Normal Distribution

Purpose: not sure which distribution? Then normal distribution!

A Useful Conclusion: $\int_{-\infty}^{\infty} e^{-x^2/c} \, dx = \sqrt{c\pi}$, where $c > 0$ is a constant.

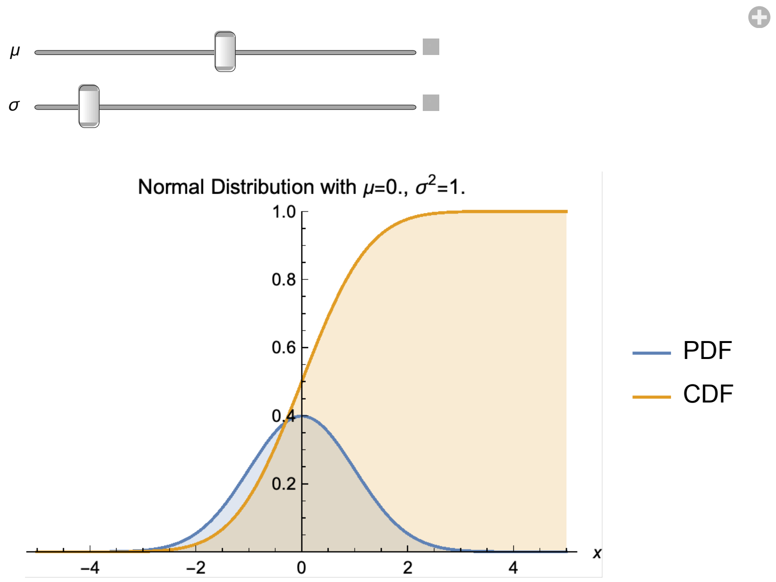
Parameter and properties:

- $\mu \in \mathbb{R}$ is the mean,
- $\sigma^2 > 0$ is the variance

Mean	Variance	PDF	CDF	MGF
μ	σ^2	$\frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$	$\frac{1}{2} \operatorname{erfc}\left(\frac{\mu - x}{\sqrt{2}\,\sigma}\right)$ where $\operatorname{erfc}(z) := 1 - \frac{2}{\sqrt{\pi}} \int_0^z e^{-t^2} \, dt$	$e^{\frac{\sigma^2 t^2}{2} + \mu t}$

- Let X be normally distributed, then $Z := \frac{X-\mu}{\sigma}$ has ***standard normal distribution*** (normal distribution with $\mu = 0$, $\sigma = 1$).

Out[ ]=



The standard normal distribution will enable you to find the value of CDF, $\Phi(x)$, simply by looking up the following table. There are many forms of standard normal table, and you can check them out here: https://en.wikipedia.org/wiki/Standard_normal_table. With the information we can calculate the following:

- $P[X \le a] = P[X < a] = \Phi(a)$
- $P[a \le X \le b] = P[X \le b] - P[X \le a] = \Phi(b) - \Phi(a)$

Out[147]=

	0.	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.	0.5	0.504	0.508	0.512	0.516	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.591	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.648	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.67	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.695	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.719	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.758	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.791	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.834	0.8365	0.8389
1.	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.877	0.879	0.881	0.883
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.898	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.937	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.975	0.9756	0.9761	0.9767
2.	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.983	0.9834	0.9838	0.9842	0.9846	0.985	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.989
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.992	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.994	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.996	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.997	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.998	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.999	0.999
3.1	0.999	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998
3.5	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998	0.9998

Estimation: From the table you can read (**how? which number in the table are you using?**)

$$P[-\sigma < X - \mu < \sigma] = 0.68$$
$$P[-2\,\sigma < X - \mu < 2\,\sigma] = 0.95$$
$$P[-3\,\sigma < X - \mu < 3\,\sigma] = 0.997$$

In[152]:=
```
CDF[NormalDistribution[0, 1], 1] // N
```

Out[152]= 0.841345

Example:

A machine in the soft drink factory produces bottles of drinks with mean $\mu = 500$ mL and standard deviation $\sigma = 1$ mL. How many mLs should a bottle have, so that 95% of the bottles produced by this machine have less drink than the bottle?

We need $Z = \frac{X-\mu}{\sigma} = \frac{X-500}{1} > 1.65 \Rightarrow X > 501.65.$

In[136]:=
```
InverseCDF[NormalDistribution[500, 1], 0.95]
```

Out[136]= 501.645

Chebyshev Inequality

Purpose: To (roughly) estimate the variance of a random variable.

Mathematical representation: for $k \in \mathbb{N} \backslash \{0\}$, $P[\,|X| \geq c] \leq \frac{E[|X|^k]}{c^k}$.

Application: $P[\,|X-\mu| \geq m\sigma] \leq \frac{1}{m^2}$.

Problems in Assignments

Two Children Paradox - Birthday Party!

What is the probability that the lady's another child is a girl given that she has a son born in July?

Solution

By Cardano's principle, we can count the number of outcomes in (G,G), (G,B), (B,G), (B,B) respectively. Each grid represent a case such as (an elder girl born in January, a younger boy born in July) and so on.

An Elder Boy

	□	□	□	□	□	□	■	□	□	□	□	□
	□	□	□	□	□	□	■	□	□	□	□	□
	□	□	□	□	□	□	■	□	□	□	□	□
	□	□	□	□	□	□	■	□	□	□	□	□
A	□	□	□	□	□	□	■	□	□	□	□	□
	□	□	□	□	□	□	■	□	□	□	□	□
Younger	■	■	■	■	■	■	■	■	■	■	■	■
Boy	□	□	□	□	□	□	■	□	□	□	□	□
	□	□	□	□	□	□	■	□	□	□	□	□
	□	□	□	□	□	□	■	□	□	□	□	□
	□	□	□	□	□	□	■	□	□	□	□	□
	□	□	□	□	□	□	■	□	□	□	□	□

(G, G)	(G, B)	(B, G)	(B, B)
Impossible	12	12	23

As result, $P[\text{another child is girl} \mid \text{boy born in July}] = \frac{12+12}{12+12+23} = \frac{24}{47}$. Actually,

$$P[\text{another child is girl} \mid \text{at least one boy}] = \frac{2}{3} \approx 0.6667$$
$$P[\text{another child is girl} \mid \text{at least one boy born in July}] = \frac{24}{47} \approx 0.5106$$
$$P\left[\text{another child is girl} \mid \text{at least one boy born on July } 1^{\text{st}}\right] = \frac{365+365}{365+365+729} \approx 0.5003$$
$$\vdots$$
$$P[\text{another child is girl} \mid \text{one elder or younger boy}] = \frac{1}{2} = 0.5$$

Conclusion: The probability is highly dependent on our observation and our ability to distinguish two events.

Problems for Discussion

Aircraft Maintenance

In order to ensure a high standard of serviceability without incurring unnecessary aircraft delays, an airline such as BEA adopts the following replacement policy for its repairable aircraft components:

- An broken component is removed from an aircraft and sent for inspection and repair.
- An immediate demand for a serviceable replacement is made to the stores and the replacement is at once fitted to the aircraft.
- When the previously broken component has been passed as serviceable, it is placed in the stores.

Suppose we know the following:

- Number of demands for a serviceable replacement is on average λ per day.
- Number of repair of a broken component is on average μ per day.

Question: how many components should the store prepare?

What is the probability of having x demands in t days?

We use the Poisson distribution with parameter λt (demands/t days).
$f_{X|t}(x) = P[X = x \mid T = t] = \frac{e^{-\lambda t}(\lambda t)^r}{x!} = \frac{e^{-\lambda t}(\lambda t)^x}{\Gamma(x+1)}$.

What is the distribution of time needed to have k components repaired?

We use the Gamma distribution with parameter $\alpha = k$, $\beta = \mu$. $f_T(t) = \frac{\mu^k}{\Gamma(k)} t^{k-1} e^{-\mu t}$. k can be interpreted as number of parallel servers that can repair the broken component.
This time interval is denoted as "lead time", the time when the shop "takes the lead" as the prepared amount is (hopefully) not used up yet, and the previous batch of repaired components has not yet come.

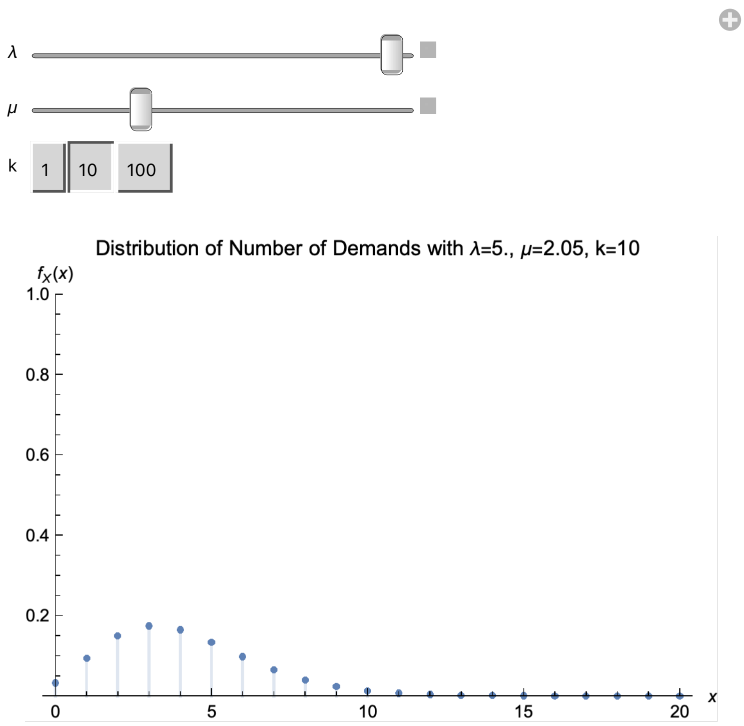
What is the probability of having x demands during the lead time?

The time interval follows a gamma distribution. Using ***marginal density*** in joint distribution (quite analogous to the ***total probability***) we can calculate

$$\begin{aligned} f_X(x) &= \int_0^\infty f_{X|t}(x)\, f_T(t)\, dt \\ &= \int_0^\infty \frac{e^{-\lambda t}(\lambda t)^x}{\Gamma(x+1)} \frac{\mu^k}{\Gamma(k)} t^{k-1} e^{-\mu t}\, dt \\ &= \frac{\mu^k \lambda^x}{\Gamma(x+1)\Gamma(k)} \int_0^\infty t^{x+k-1} e^{-(\mu+\lambda) t}\, dt \\ &= \frac{\mu^k \lambda^x}{(\mu+\lambda)^{x+k}\Gamma(x+1)\Gamma(k)} \int_0^\infty z^{x+k-1} e^{-z}\, dz \qquad (z := (\mu+\lambda)\, t) \\ &= \frac{\Gamma(x+k)}{\Gamma(x+1)\Gamma(k)} \frac{\mu^k \lambda^x}{(\mu+\lambda)^{x+k}} = \binom{k+x-1}{k-1} \left(\frac{\lambda/\mu}{1+\lambda/\mu}\right)^r \left(\frac{1}{1+\lambda/\mu}\right)^k \end{aligned}$$

So, the probability of having x demands follows a **negative binomial distribution** with parameters $p = \frac{\lambda/\mu}{1+\lambda/\mu}$ and $r = k$. This shows that neg. binomial distribution can be seen as a combination of Poisson and gamma distribution.

Out[130]=



Now, suppose we have n components prepared at the store at any time. The probability of shortage of component will be ***risk level*** := P[having more than $n+1$ demands] $= P[x \geq n+1] = \sum_{x=n+1}^{\infty} f_X(x)$. Different number of k can be used to estimate the risk level of having n components prepared at the store. The rest of the solution will be posted through Piazza.