$X$ is a r.v. with

\[
F(x) = \begin{cases} 
1 - e^{-\lambda x} & x \geq 0 \\
0 & x < 0 
\end{cases}
\]

\[
f(x) = \lambda e^{-\lambda x}
\]

\[
\frac{d}{dx} F \rightarrow f
\]

"Random - Nixodyne"\n
\[
P(X = x) = 0 \text{ for } X \text{ in }
\]

\[
F_X(x) = P(X \leq x)
\]

\[
S_X(x) = P(X > x)
\]
\[ S_n(x) = \begin{cases} 0 & \text{if } x < X_{(1)} \\ \frac{i}{n} & \text{if } X_{(i)} \leq x < X_{(i+1)} \\ 1 & \text{if } x \geq X_{(n)} \end{cases} \]

\( S_n(x) \) is an estimator of \( F_n(x) \)

"EDF is a step-fn."

"It jumps only at the distinct ordered values." → "Distinct ordered values," → "Height of the jump = \( k \frac{S_n(x)}{n} \) of values tied at \( X_{(j)} \)"
Define $T_n(z) = n \cdot S_n(z)$

$T_n(z) = \text{Total \# of sample values } \leq \text{some specified } z$

Theorem: For any real $z$, the r.v $T_n(z)$

Define $\delta_i(z) = \{ 1 \text{, if } X_i \leq z \}$

$\text{Indicator function}\tries{\text{for } \delta_i(z)} = E(\delta_i(z)) = \Pr(\text{Event})$
\[ E(\delta_i(x)) = \mathbb{P}(X_i \leq x) + 0. \mathbb{P}(X_i > x) \]
\[ = \mathbb{P}(X \leq x) \]
\[ = F_X(x) \]

See that, \( \delta_1(x), \ldots, \delta_n(x) \) are i.i.d. Bernoulli \((\theta)\)

\[ \Theta = \mathbb{P}(\delta_i(x) = 1) = E(\delta_i(x)) \]

\[ T_\Theta(x) = n \Sigma \delta_i(x) \]

\( T_\Theta(x) \) is the sum of \( n \) i.i.d. \( \text{Ber}(\theta) \)
\[ X_i, \ i = 1(1) \sim N(\mu_i, \sigma_i^2) \]

\[ X, \ \iids \sim N(\mu, \sigma^2) \]

& \[ X, \ldots, X_n \text{ are i} \]

\[ \mu_i = \mu, = \cdot = \mu_n = \mu \]

\[ X, \ldots, X_n \sim \text{Bin}(\theta) \]

\[ T = \sum X_i \sim \text{Bin}(n, \theta) \]

\[ T = \sum \frac{X_i}{\theta} \sim \text{Bin}(n, \theta) \]
\[ T_n \sim Bin(n, \theta = F_X(x)) \]

1. \[ E(S_n(x)) = F_X(x) \]

\[ V_S(S_n(x)) = \frac{F_X(x)(1 - F_X(x))}{n} \]

"We think \( S_n(x) \) should approximate \( F_X(x) \)."
$S_n(x)$ "converges uniformly" to $F_X(x)$ with probability 1.

& this is written as

$$P \left[ \lim_{n \to \infty} \sup_{-\infty < x < \infty} |S_n(x) - F_X(x)| = 0 \right] = 1$$

"Gleivenko–Cantelli Theorem"
Converging with prob. 1
"almost sure convergence"
"g.s.e.t. convergence"
"in distribution"

\[ F_i \xrightarrow{d} F \quad \text{Excellent!!} \]
No worries
\[
\lim_{n \to \infty} P \left\{ \frac{\sqrt{n} \left( S_n(x) - F_x(x) \right)}{\sqrt{\frac{F_x(x)}{1-F_x(x)}}} \leq \epsilon \right\} \xrightarrow{\epsilon} N(0,1) \\
\text{CLT (for } S_n(x) \text{) = } \Phi(\epsilon)
\]
To find the c.d.f of $X^{(r)}$:

$F_{X^{(r)}}(x) = P(X^{(r)} \leq x)$

$= \sum_{i=r}^{n} \binom{n}{i} \left[ F_X(x) \right]^i \left(1 - F_X(x) \right)^{n-i}$
\[ n \, S_n(x) \sim \text{Bin}(n, F_X(x)) \]

\(X(n) \leq t\) happens, "iff"

(if & only if) at least \(\gamma\) of the \(X\)'s \(\leq t\).
Then, what is the p.d.f of $X(r)$?

$F_X(x) = f_x(x)$ exists.

$\therefore f_X(x) = \frac{n!}{(r-1)! (n-r)!} \left[ F_X(x) \right]^{r-1} \left[ 1 - F_X(x) \right]^{n-r} \int_{F_X(x)}^{1} f_x(z) dz$

$n$ total pts.
Probability Integral Transform (PIT)

Let $X$ be a r.v. with c.d.f $F_X(x)$

$F_X(\cdot)$ is continuous.

Then, $Y = F_X^{-1}(U) \sim U(0,1)$

$\text{VV FUP} - \text{Result}$.
This means, if \( X_1, X_2, \ldots, X_n \) is a random sample from a population with c.d.f. \( F_X(x) \) then \( F_X(X_1), F_X(X_2), \ldots, F_X(X_n) \) is a random sample from the Uniform distribution.
PIT Proof excluded!!

Ex: Let \( X \sim \text{Exp}(\text{Mean} = 2) \)

\[ c.d.f \ of \ X \ is \ F_X(x) = 1 - e^{-\frac{x}{2}} \]

Then by Pitman's Th, the s.v.

\[ Y = f_X(x) \sim U(0, 1) \]

\[ 1 - e^{-\frac{x}{2}} = u \]
Solve for $n$?

$$X = -2 \cdot \ln (1-u)$$

This gives me a way to generate some $\exp(2)$ r.v.

In R:

```r
u <- runif(0,1)
X.exp <- -2 * log(1-u)  # a random exp(2). r.v.
```
H. W. 1000

1. Generate Exp (Mean = 10) samples using
   (i) P/T Method
   (ii) Using \( \alpha \cdot \exp \) 

& compare !!
Uniform(0, 1) \[ ? \]
Rectangular( a, b ) \[ \leq \] W
Conservative Tests

We are comparing 2 brands of tires

$H_0: \mu_1 - \mu_2 = 0$, $H_A: \mu_1 - \mu_2 \neq 0$

$t_1 = 23.58$, $s_1 = 3100$

$t_2 = 25.86$, $s_2 = 3600$
\[ t - \text{test} : \]
\[ t = \frac{(\bar{x}_1 - \bar{x}_2) - (\mu_1 - \mu_2)}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \]
\[ s_p = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}} \]

We get \( t = -1.663 \)
We can focus on either a $t$-table or a $z$-table.

In R? $p$-value?

> 2 * pbinom(-1.663, 22)  # 0.0963

> 2 * pt(-1.663)  # 0.11049  # x

> 2 * pt(1.663)  # 2
A conservative test is a test where it doesn't reject the NULL often enough (i.e., a test having a stated level of significance > true level).

It can be determined by simulation.
Set $d = 0.05$; consider 2 tests $T_1$ & $T_2$

Do 10000 runs/repetitions

If you get 400 rejections, $(T_1)$
   You have a Conservative Test $T_1$.

If you get 600 rejections:
   You have a Non-conservative Test $T_2$. 
   (anti-...