statistics cheat sheet

Nasser M. Abbasi

Sumemr 2008 page compiled on July 2, 2015 at 6:07pm

problem: phone calls received at rate $\lambda = 2$ per hr. If person wants to take 10 min shower, what is probability a phone will ring during that time?

answer: first change to $\omega = \lambda \frac{10}{60} = 2\frac{10}{60} = .3333$, now we want $P(X \geq 1) = 1 - P(X \leq 1) = 1 - P(0)$

but $P(k) = \frac{\lambda^k}{k!} e^{-\lambda}$, but remember, we are using ω, so $P(k) = \frac{\omega^k}{k!} e^{-\omega}$ so $P(0) = \frac{.3333^0}{0!} e^{-.3333} = 0.777$

so $P(X \geq 1) = 1 - .777 = 0.283$, so 28% change the phone will ring.

How long can shower be if they wish probability of receiving no phone calls to be at most 0.5?

$P(0) = 0.5 = \frac{\omega 0}{0!} e^{-\omega} \rightarrow 0.5 = e^{-\omega}$ hence $\ln 0.5 = -\omega \rightarrow \omega = 0.693$, so $\lambda \frac{x}{60} = 0.693 \rightarrow x = 20.7$ minutes

To find quantile, say $\frac{1}{4}$, first find an expression for $F(x)$ as function of x, then solve for x in $F(x) = .25$

For median, solve for x in $F(x) = .5$

properties of CDF: 1. Show $F(x) \geq 0$ for all x. Do this by showing $F'(x) \geq 0$, and show limit $F(x) \rightarrow 1$ as $x \rightarrow \infty$ and limit $F(x) \rightarrow 0$ as $x \rightarrow -\infty$. And $\boxed{P(k_1 \leq T < k_2) = F(k_2) - F(k_1)}$

properties of pdf:

1. piecewise continuous
2. pdf$(x) \geq 0$
3. $\int_{-\infty}^{\infty} pdf(x) = 1$

remember $\frac{d}{dx} \tan^{-1} x = \frac{1}{1+x^2}$

The geometric distribution is the only discrete memoryless random distribution. It is a discrete analog of the exponential distribution. continuous

Some relations

$$\sum_{k=1}^{n} k = \frac{1}{2} n(n+1)$$

Geometric sum

$$\sum_{k=0}^{n} r^k = \frac{1 - r^{n+1}}{1 - r}$$

if $-1 < r < 1$, then

$$\sum_{k=0}^{\infty} r^k = \frac{1}{1 - r}$$

if the sum is from 1 then

$$\sum_{k=1}^{n} r^k = \frac{r\left(1 - r^{n+1}\right)}{1 - r}$$

if $-1 < r < 1$, then

$$\sum_{k=1}^{\infty} r^k = \frac{r}{1 - r}$$

$$\Gamma(x) = \int_0^{\infty} u^{x-1} e^{-u} du$$

$$\Gamma(n) = (n-1)!$$

$$\frac{d}{dx}\ln(x) = \frac{1}{x}$$

$$\int \ln(y)\,dy = -y + y\ln(y)$$

$$\int \frac{1}{y}dy = \ln(y)$$

$$\begin{pmatrix} & n & \\ n_1 & n_2 & n_3 \end{pmatrix} = \frac{n!}{n_1!\; n_2!\; n_3!}$$

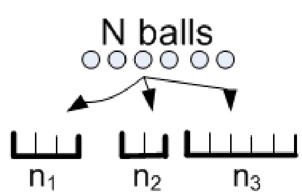
If given joint density $f_{XY}(x,y)$ and asked to find conditional $P(X|Y) = \frac{f_{XY}(x,y)}{f_Y(y)}$ so need to find marginals. Marginal is found from $f_Y(y) = \int_x f_{XY}(x,y)\,dx$, and $f_X(x) = \int_y f_{XY}(x,y)\,dy$

To convert from x, y to polar, example: given $f(x,y) = c\sqrt{1-(x^2+y^2)}$ find c, where $x^2+y^2 \le 1$, then write

$$c\int_{\theta=-\pi}^{\theta=\pi}\int_{r=0}^{r=1}\sqrt{1-r^2}\,r\,dr\,d\theta$$

Use identity above.

law of total probablity: if we know $Y|X$ and X and want to know distribution of Y, then $f(Y) = \int_{-\infty}^{\infty} f_{Y|X}(y|x)\,f_X(x)\,dx$



The Number of WAYS to distribute n balls into 3 buckets of different sizes is given by the multinomial $\begin{pmatrix} & n & \\ n_1 & n_2 & n_3 \end{pmatrix} = \frac{n!}{n_1!\,n_2!\,n_3!}$

In[6]:= $\int \sqrt{a^2 - x^2}\,dx$

Out[6]= $\frac{1}{2}\left(x\sqrt{a^2-x^2} + a^2\,\mathrm{ArcTan}\left[\frac{x}{\sqrt{a^2-x^2}}\right]\right)$

In[5]:= $\int \sqrt{a^2 + x^2}\,dx$

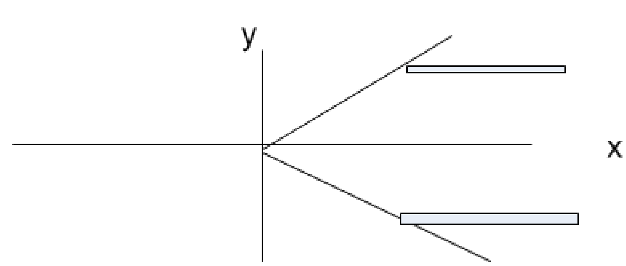
Out[5]= $\frac{1}{2}x\sqrt{a^2+x^2} + \frac{1}{2}a^2\,\mathrm{Log}\left[x + \sqrt{a^2+x^2}\right]$

$r = \int (x + a\,y)^n\,dy;$

Out[29]= $\frac{(x + a\,y)^{1+n}}{a\,(1+n)}$

In[2]:= $\int r\sqrt{1-r^2}\,dr$

Out[2]= $-\frac{1}{3}\left(1-r^2\right)^{3/2}$



If we need to find marginal density f_Y, then do it in 2 parts. For $y > 0, f_Y(y) = \int_{x=y}^{x=\infty} f(x,y)dx$ and For $y < 0, f_Y(y) = \int_{x=-y}^{x=\infty} f(x,y)dx$ and leave it at that. do not Add them.

$$Z = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} \to N(0,1)$$

$$T = \frac{\bar{X}_n - \mu}{S_n/\sqrt{n}} \to T(n)$$

where S_n is *std* of the sample.

Note Var(sample) has chi square (n) distribution.

CI for T:

$$\Pr\left(-A < \frac{\bar{X}_n - \mu}{S_n/\sqrt{n}} < A\right) = 1 - \alpha$$

$$\Pr\left(\bar{X}_n - A\frac{S_n}{\sqrt{n}} < \mu < \bar{X}_n + A\frac{S_n}{\sqrt{n}}\right) = 1 - \alpha$$