1 Discrete Distributions

1.1 Bernoulli and Binomial

The Bernoulli distribution is the simplest case of the Binomial distribution, where we only have one trial ($n = 1$). Let us say that X is distributed Bern(p). We know the following: A trial is performed with probability p of "success", and X is the indicator of success: 1 means success, 0 means failure.

Let us say that X is distributed Bin(n, p). We know the following: X is the number of "successes" that we will achieve in n independent trials, where each trial is either a success or a failure, each with the same probability p of success. We can also write X as a sum of multiple independent Bern(p) random variables. Let $X \sim \text{Bin}(n, p)$ and $X_j \sim \text{Bern}(p)$, where all of the Bernoullis are independent. Then $X = X_1 + X_2 + \ldots + X_n$.

1. PMF, MGF, mean and variance of $X \sim \text{Binom}(n, p)$
   (a) **PMF**:
   $$f(x) = \binom{n}{x} p^x (1-p)^{n-x},$$
   $$x = 0, 1, ...n$$
   $$0 < p < 1$$
   (b) **MGF**: $M_X(t) = [(1-p) + pe^t]^n$
   (c) **Mean and Variance**:
   $$E[X] = np, \; Var(X) = np(1-p)$$

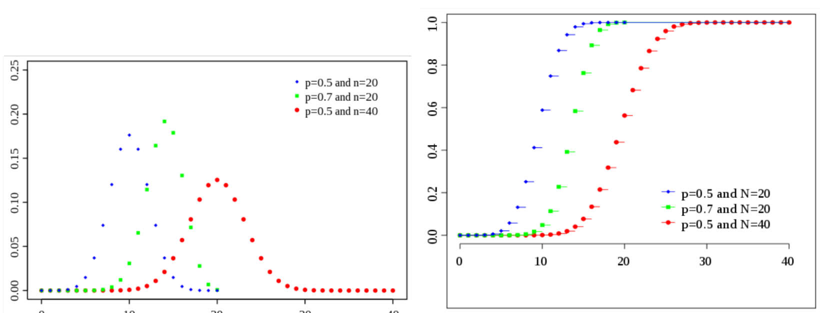

Figure 1.1: Bernoulli PMF (left) CDF (right)

2. **Additive property**
   If for $i = 1, 2, ..., k$, $X_i \sim \text{Binom}(n_i, p)$, then $\sum_{i=1}^{k} X_i \sim \text{Binom}(\sum_{i=1}^{k} n_i, p)$
3. Random sample $X_1, ..., X_n \sim \text{Ber}(p)$ where p is **target parameter**: exponential family? sufficient statistic? minimal sufficient statistic? complete statistic? Fisher information? UMVUE?
   (a) **Exponential family** form: $(1-p)^n \exp\left(\sum_{i=1}^{n} x_i log \frac{p}{1-p}\right)$
   (b) **Minimal sufficient and complete statistic**: $\sum_{i=1}^{n} X_i$
   (c) **Fisher information**: $\frac{n}{p(1-p)}$
   (d) **UMVUE**: $\frac{1}{n} \sum_{i=1}^{n} X_i$
   (e) **MLE** $\hat{p} = \frac{1}{n} \sum_{i=1}^{n} X_i$
4. **Conjugate prior** of p?
   $X|p \sim \text{Binom}(n, p)$
   $p \sim \text{beta}(\alpha, \beta)$.
   Posterior distribution $p|(X = x) \sim \text{beta}(\alpha + \sum_{i=1}^{n} x_i, n + \beta - \sum_{i=1}^{n} x_i)$

5. **Related Distributions**

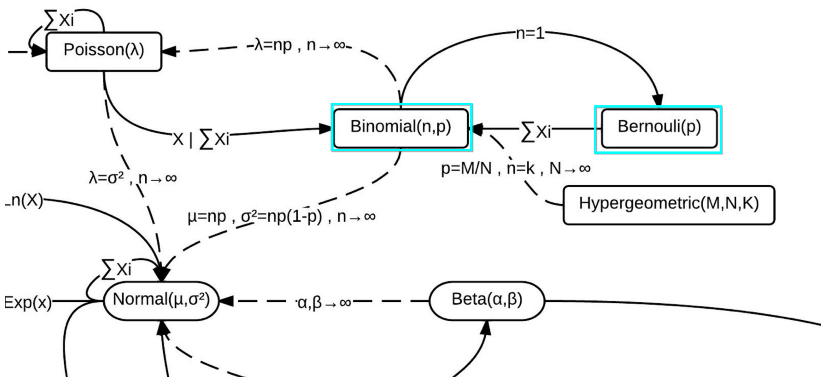

Figure 1.2: Bernoulli and Binomial related distributions

(a) If $X \sim \text{Ber}(p)$ then $\sum_{i=1}^{n} X_i \sim \text{Bin}(n, p)$

(b) **Poisson and Normal Approximations**

i. Poisson Approximation (Casella example 2.3.13): for large n and small np, $\text{Binom}(n, p) \xrightarrow{d} Pois(\lambda)$,where $\lambda = np$.

ii. Normal Approximation (via CLT): $\frac{X}{n}$ follows approximate Normal distribution with mean p and variance $\frac{p(1-p)}{n}$.

(c) **Indicator Function** - $I_{x \in A}(x) \sim \text{Bern}(p)$ where $p = P(x \in A)$, and sum of n i.i.d indicators (support A) ~Bin(n, p).

6. Example problems and key steps

1.2 Poisson

Let us say that X is distributed Pois(X). We know the following: There are rare events (low probability events) that occur many different ways (high possibilities of occurences) at an average rate of λ occurrences per unit space or time. The number of events that occur in that unit of space or time is X.

1. **PMF, MGF, mean and variance** of $X \sim \text{Pois}(\lambda)$

   (a) **PMF**:
   $$f(x) = \frac{\lambda^x e^{-\lambda}}{x!}$$
   $$x = 0, 1, ...$$
   $$\lambda > 0$$

   (b) **MGF**: $M_X(t) = \exp(\lambda(e^t - 1))$

   (c) **Mean and Variance**:
   $$E[X] = \lambda,\ Var(X) = \lambda$$

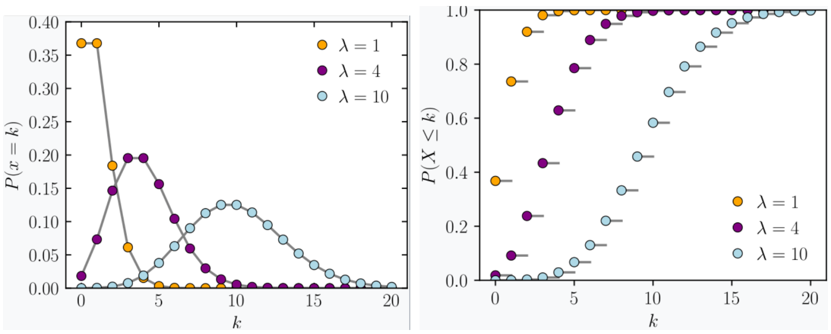

Figure 1.3: Poisson PMF (left) CDF (right)

2. Random sample $X_1, ..., X_n \sim \text{Pois}(\lambda)$ where λ is target parameter: exponential family? sufficient statistic? minimal sufficient statistic? complete statistic? Fisher information? UMVUE?

   (a) **Exponential family** form: $\prod_{i=1}^{n} [x_i!]^{-1} \cdot e^{-n\lambda} \exp\left(\sum_{i=1}^n x_i \log \lambda\right)$

   (b) **Minimal sufficient and complete statistic**: $\sum_{i=1}^n X_i$

   (c) **Fisher information**: $\frac{n}{\lambda}$

   (d) **UMVUE**: $\frac{1}{n}\sum_{i=1}^n X_i$

   (e) **MLE** $\hat{\lambda} = \frac{1}{n}\sum_{i=1}^n X_i$

3. **Conjugate prior** of λ?

   $X|\lambda \sim \text{Pois}(\lambda)$

   $\lambda \sim \text{Gamma}(\alpha, \beta)$

   Posterior distribution $\lambda|(X = x) \sim \text{Gamma}(\alpha + \sum_{i=1}^n x_i, \frac{1}{n+\frac{1}{\beta}})$

   Note, above is if Gamma *pdf* is written as $f(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-\frac{x}{\beta}}$

   If Gamma *pdf* is written as $f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-x\beta}$ then posterior distribution is written as $\lambda|(X = x) \sim \text{Gamma}(\alpha + \sum_{i=1}^n x_i, n + \beta)$

4. **Related Distributions**

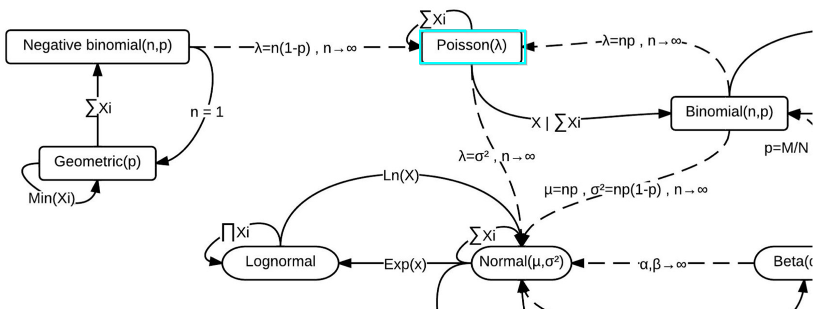

Figure 1.4: Poisson and related distributions

(a) **Additive property**

i. For $X_i \overset{ind}{\sim} \text{Pois}(\lambda_i)$, then $\sum_{i=1}^{k} X_i \sim \text{Pois}(\sum_{i=1}^{k} \lambda_i)$

(b) **Relation to Binomial** distribution

i. If $X \sim \text{Pois}(\lambda_x)$ and $Y \sim \text{Pois}(\lambda_y)$ and they are independent, then $X|X+Y=n \sim \text{Binom}\left(n, \frac{\lambda_x}{\lambda_x+\lambda_y}\right)$

(c) **Normal Approximation** (via CLT)

i. If $X \sim \text{Pois}(\lambda)$ then $Z = \frac{X-\lambda}{\sqrt{\lambda}} \xrightarrow{d} N(0,1)$ for $n \to \infty$

5. Example problems and key steps

1.3 Negative Binomial

Let us say that X is distributed NBin(r, p). We know the following: X is the number of "failures" that we will have before we achieve our rth success. Our successes have probability p.

1. **PMF, MGF, mean and variance** of $X \sim \text{NegBinom}(r, p)$
   (a) **PMF**:
   $$f(x) = \binom{x+r-1}{x} p^r (1-p)^x$$
   $$x = 0, 1, ...n$$
   $$0 < p < 1$$
   (b) **MGF**: $M_X(t) = \left(\frac{p}{1-(1-p)e^t}\right)^r$, $t < -\log p$
   (c) **Mean and Variance**:
   $$E[X] = \frac{r(1-p)}{p},\ Var(X) = \frac{r(1-p)}{p^2}$$

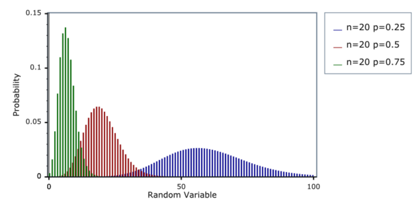

Figure 1.5: Negative Binomial PMF

2. **Additive property**
   If for $i = 1, 2, ..., k$, $X_i \overset{ind}{\sim} \text{NegBinom}(r_i, p)$, then $\sum_{i=1}^k X_i \sim \text{NegBinom}(\sum_{i=1}^k r_i, p)$
3. Random sample $X_1, ..., X_n \overset{i.i.d}{\sim} \text{NegBinom}(r, p)$ where p is target parameter and r is known: exponential family? sufficient statistic? minimal sufficient statistic? complete statistic? Fisher information? UMVUE?
   (a) **Exponential family** form: $\prod_{i=1}^n \binom{x_i+r-1}{x_i} \cdot p^{nr} \exp\left(\sum_{i=1}^n x_i \log(1-p)\right)$
   (b) **Minimal sufficient and complete statistic**: $\sum_{i=1}^n X_i$
   (c) **Fisher information**: $\frac{nr}{p^2(1-p)}$
   (d) **UMVUE**:
      i. $\frac{1-p}{p}$: $\frac{\sum_{i=1}^n x_i}{nr}$
      ii. p^r: $P(X_i = 0)$, $E(X_i = 0|\sum_{i=1}^n X_i)$, $E(\sum_{i=1}^n x_i) = \frac{nr(1-p)}{p}$
      $T = I_{(X_i=1)}(X_i)$ where $E(T) = P(X_i = 1) = p^r$. Let $S = \sum_{i=1}^n X_i$ and consider $E[T|S]$ (Rao Blackwell).
      $$E[T|S] = E[T = 1|S = s] = \frac{p^t \binom{s-t-1}{r-t-1} p^{r-t}(1-p)^{s-r}}{\binom{s-1}{r-1} p^r (1-p)^{s-r}} = \frac{(s-t-1)!(r-1)!}{(s-1)!(r-t-1)!} \text{ is UMVUE}$$
4. **Conjugate prior** of p?
   $X|p \sim \text{NegBinom}(r, p)$
   $p \sim \text{beta}(\alpha, \beta)$.
   Posterior distribution $p|(X = x) \sim \text{beta}(\alpha + nr, \beta + \sum_{i=1}^n x_i)$

5. **Related Distributions**

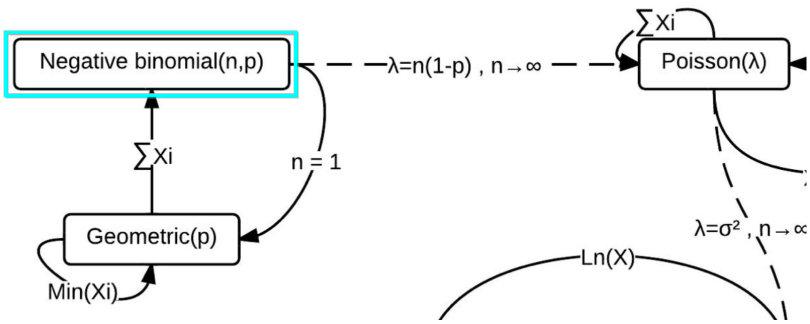

Figure 1.6: Negative Binomial and related distributions

(e) If $X \sim \text{NegBinom}(r, p)$ for $r = 1$, it becomes geometric distribution with succss probability p

(f) With $p = \frac{r}{r+\lambda}$, with $X \sim \text{NegBinom}(r, \frac{r}{r+\lambda})$, then $X \xrightarrow{d} \text{Pois}(\lambda)$ for $r \to \infty$.

6. **Example problems** and key steps

1.4 Geometric***

Let us say that X is distributed Geom(p). We know the following: X is the number of "failures" that we will achieve before we achieve our first success. Our successes have probability p.

Number of Bernoulli trials before a success. (Negative binomial with $r = 1$)

1. **PMF, MGF, mean and variance** of $X \sim \text{Geom}(p)$

   (a) **PMF**:

   i. x trials before first success:
   $$f(x) = p(1-p)^{x-1}$$
   $$x = 1, 2, ...$$
   $$0 < p < 1$$

   ii. x failures before first success:
   $$f(x) = p(1-p)^{x}$$
   $$x = 0, 1, 2, ...;$$
   $$0 < p < 1$$

   (b) **CDF**:

   i. x trials before first success: $F(t) = 1 - (1-p)^t,\ t = 1, 2, ...$

   ii. x failures before first success: $F(t) = 1 - (1-p)^{t+1},\ t = 0, 1, 2, ...$

   (c) **MGF**: $M_X(t) = \left(\frac{pe^t}{1-(1-p)e^t}\right)^r$, $t < -\log(1-p)$

   (d) **Mean and Variance**:
   $$E[X] = \frac{1}{p},\ Var(X) = \frac{(1-p)}{p^2} \text{ for } f(x) = p(1-p)^{x-1},\ x = 1, 2, ...$$
   $$E[X] = \frac{1-p}{p},\ Var(X) = \frac{(1-p)}{p^2} \text{ for } f(x) = p(1-p)^{x},\ x = 0, 1, 2, ...$$

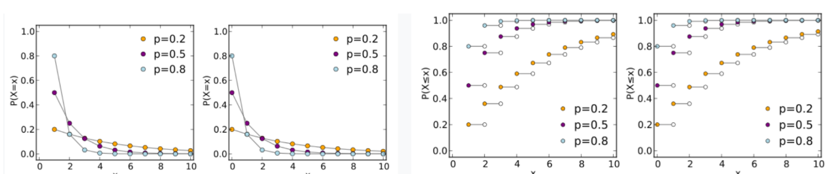

Figure 1.7: Geometric PMF (left) CDF (right)

2. Random sample $X_1, ..., X_n \overset{i.i.d}{\sim} \text{Geom}(p)$ where p is target parameter: exponential family? sufficient statistic? minimal sufficient statistic? complete statistic? Fisher information? UMVUE?

   (a) **Exponential family** form: $\left(\frac{p}{1-p}\right)^n \exp\left(\sum_{i=1}^n x_i \log(1-p)\right)$ for $f(x) = p(1-p)^{x-1},\ x = 1, 2, ...$

   (b) **Minimal sufficient and complete statistic**: $\sum_{i=1}^n X_i$

   (c) **Fisher information**: $\frac{n}{p^2(1-p)}$

   (d) **UMVUE**:

   i. $\frac{1}{p}$: $\frac{\sum_{i=1}^n x_i}{n}$ for $f(x) = p(1-p)^{x-1},\ x = 1, 2, ...$

   ii. $\frac{1-p}{p}$: $\frac{\sum_{i=1}^n x_i}{n}$ for $f(x) = p(1-p)^{x},\ x = 0, 1, 2, ...$

   i. p: $\sum_{i=1}^n x_i \sim \text{NegBin}(n, p)$ is sufficient and complete, by L-S UMVUE is $\dfrac{(n-1)}{(n-1+\sum_{i=1}^n x_i)}$
   $E\left(\frac{(n-1)}{(n-1+n\bar{X})}\right) = \sum_{k=0}^{\infty} \frac{(n-1)}{(n-1+k)} \binom{n+k-1}{k} p^n (1-p)^k = \sum_{k=0}^{\infty} \frac{(n-1)}{(n-1+k)} \frac{(n+k-1)!}{(n-1)!k!} p^n (1-p)^k$
   $\sum_{k=0}^{\infty} \frac{(n+k-2)!}{(n-2)!k!} p^n (1-p)^k = p \sum_{k=0}^{\infty} \binom{n+k-2}{k} p^{n-1} (1-p)^k = p$

3. **Conjugate prior** of p?

   $X|p \sim \text{Geom}(p)$

   $p \sim \text{beta}(\alpha, \beta)$.

   Posterior distribution $p|(X = x) \sim \text{beta}(\alpha + n, \beta + \sum_{i=1}^{n} x_i - n)$

4. **Related Distributions**

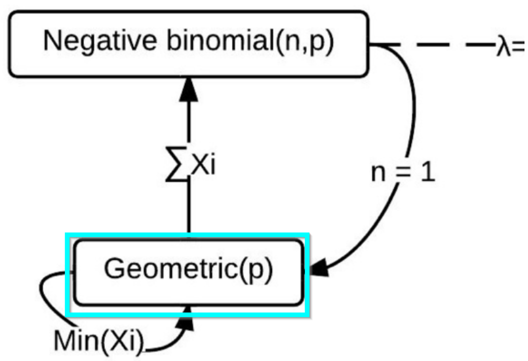

Figure 1.8: Geometric and related distributions

   (a) If $X \sim \text{Geom}(p)$ then $\sum_{i=1}^{r} X_i \sim \text{NegBinom}(r, p)$

   (b) If $X \sim \text{Geom}(p_1)$ and $Y \sim \text{Geom}(p_2)$ are independent, and $Z = \min(X, Y)$, then $Z \sim \text{Geom}(1 - [1 - p_1][1 - p_2])$

5. **Example problems** and key steps

1.5 Hypergeometric

In a population of M desired objects and N undesired objects, x is the number of "successes" we will have in a draw of K objects, without replacement. The draw of K objects is assumed to be a simple random sample (all sets of K objects are equally likely).

1. **PMF, MGF, mean and variance** of $X \sim \text{Hypergeometric}(N, M, K)$

   (a) **PMF**:

   $$P(X = x|N, M, K) = \frac{\binom{M}{x}\binom{N-M}{K-x}}{\binom{N}{K}}$$

   $$x = 0, 1, ..., K \text{ and } \max(0, K + M - N) \leq x \leq \min(M, K)$$

   $$M \in \{0, 1, 2, ..., N\}, K \in \{0, 1, 2, ..., N\},\ N \in \{0, 1, 2...\}$$

   (b) **MGF**: $M_X(t) = \frac{\binom{N-M}{K} {}_2F_1(-K,-M;N-M-K+1;e^t)}{\binom{N}{K}}$

   (c) **Mean and Variance**:

   $$E[X] = \frac{KM}{N},\ Var(X) = \frac{KM}{N}\left(\frac{(N-M)(N-K)}{N(N-1)}\right)$$

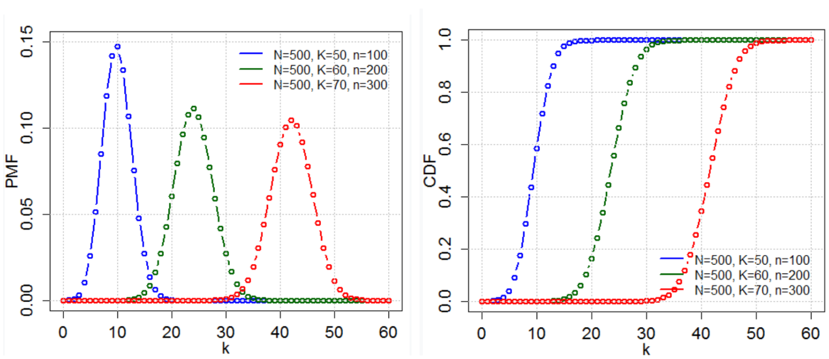

Figure 1.9: Hypergeometric PMF (left) CDF (right)

2. Random sample $X_1, ..., X_n \overset{i.i.d}{\sim} H\text{Geom}(N, M, K)$ where p is target parameter: exponential family? sufficient statistic? minimal sufficient statistic? complete statistic? Fisher information? UMVUE?**

   (a) Exponential family form:

   (b) Minimal sufficient and complete statistic: **

   (c) Fisher information: **

   (d) UMVUE:

3. **Conjugate prior** of p?

   $X|N, M, K \sim \text{HGeom}(N, M, K)$

   $p \sim \text{BetaBin}(n, \alpha, \beta)$.

   Posterior distribution $p|(X = x) \sim \text{BetaBin}(?)$

4. **Related Distributions**

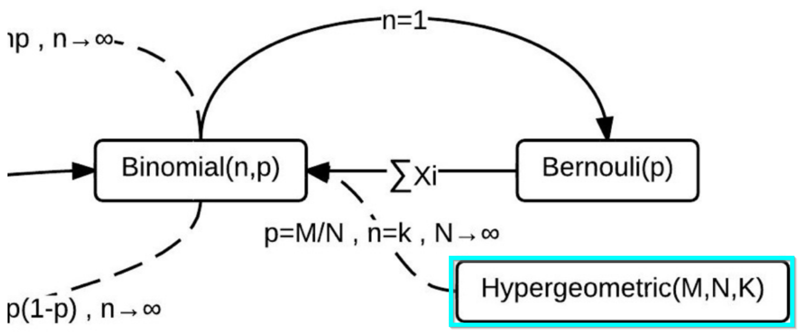

Figure 1.10: Hypergeometric and related distributions

If $X \sim \text{Hypergeometric}(N, M, K)$ and $p = \frac{M}{N}$

(a) If $K = 1$ then $X \sim \text{Ber}(p = \frac{M}{N})$

(b) If $Y \sim \text{Bin}(n, p)$ then Y models the number of successes in the analagous sampling problem *with* replacement. If N and M are large compared to K, and p is not close to 0 or 1, then X and Y have similar distributions, i.e. $P(X \leq c) \approx P(Y \leq c)$.

(c) If K is large and M and N are large compared to K, and p is not close to 0 or 1, then $P(X \leq x) \approx \Phi\left(\frac{x-Kp}{\sqrt{Kp(1-p)}}\right)$

5. **Example problems** and key steps

2 Continuous Distributions

2.1 Continuous Uniform

Let us say that U is distributed Unif(a, b). We know the following: **Properties of the Uniform** For a Uniform distribution, the probability of a draw from any interval within the support is proportional to the length of the interval.

1. **PDF, CDF, MGF, mean and variance** of $X \sim$ Unif(a, b)

   (a) **PDF**:
   $$f(x) = \frac{1}{b-a}$$
   $$a \leq x \leq b$$
   $$a \in \mathbb{R}, b \in \mathbb{R}$$

   (b) **CDF**: $f(t) = \frac{x-a}{b-a}$
   $a \leq t \leq b$
   $a \in \mathbb{R}, b \in \mathbb{R}$

   (c) **MGF**: $M_X(t) = \frac{e^{bt}-e^{at}}{(b-a)t}$

   (d) **Mean and Variance**:
   $$E[X] = \frac{a+b}{2}, \; Var(X) = \frac{(b-a)^2}{12}$$

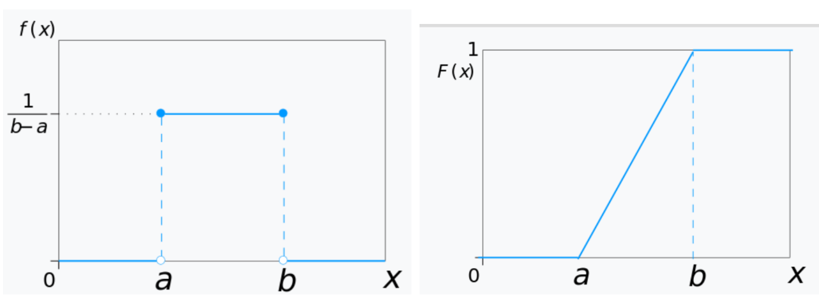

Figure 2.1: Uniform PDF (left) CDF (right)

2. Random sample $X_1, ..., X_n \sim$ Unif$(0, b)$ where b is target parameter: exponential family? sufficient statistic? minimal sufficient statistic? complete statistic? Fisher information? UMVUE?

   (a) **Not exponential family**

   (b) **Scale family**: the standard distribution is Unif$(0, 1)$ and $\frac{X_i}{b} \sim$ Unif$(0, 1)$ (b is the scale parameter)

   (c) **Minimal sufficient and complete statistic**: $X_{(n)}$

   (d) **Fisher information**: $\frac{n}{b^2}$

   (e) **UMVUE**: $\frac{n+1}{n} X_{(n)}$ whose variance is $\frac{b^2}{n(n+2)}$ less than CRLB, indicating C-R inequality is not applicable for this population. (read Casella example 7.3.13)

3. Random sample $X_1, ..., X_n \sim$ Unif(a, b) where a, b are target parameters: exponential family? sufficient statistic? minimal sufficient statistic? complete statistic? Fisher information? UMVUE?

   (f) **Not exponential family**

   (g) **Scale family**: the standard distribution is Unif$(0, 1)$ and $\frac{X_i - a}{b-a} \sim$ Unif$(0, 1)$ (a is the location parameter and $b - a$ is the scale parameter)

(h) **Minimal sufficient and complete statistic**: $(X_{(1)}, X_{(n)})$.

(i) **MLE** of a is $X_{(1)}$ and **MLE** of b is $X_{(n)}$. **MLE** of $\theta = b - a$ is $\hat{\theta} = X_{(n)} - X_{(1)}$ by invariance property of MLE.

4. **Related Distributions**

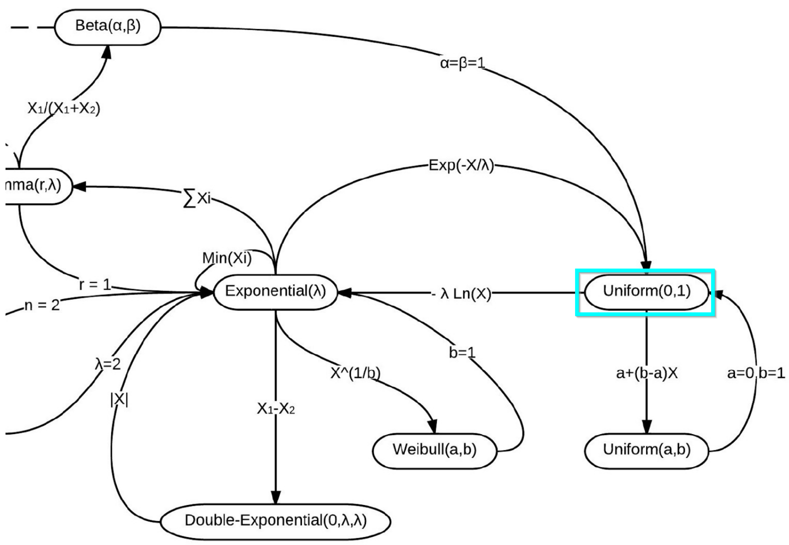

Figure 2.2: Uniform related distributions

(j) If $X \sim \text{Unif}(a,b)$ then single order statistics (Casella example 5.4.5): $\frac{X_{(j)}-a}{b-a} \sim \text{Beta}(j, n-j+1)$*

(k) If $X \sim \text{Unif}(0,1)$, then $-\lambda \log X \sim \text{EXP}(\lambda)$ (λ is scale parameter)

(l) Bivariate order statistics (assuming $i < j$): $\left(\frac{X_{(i)}-a}{b-a}, \frac{X_{(j)}-a}{b-a}\right) \sim \text{Dir}(i, j-i, n-j+1)$ (Casella exercise 4.40 for Dirichlet distribution)

5. **Example problems** and key steps

2.2 Gamma

Let us say that X is distributed Gamma(a, β). We know the following: You sit waiting for shooting stars, where the waiting time for a star is distributed EXP(β). You want to see n shooting stars before you go home. The total waiting time for the nth shooting star is Gamma(n, β).

1. **PDF, MGF, mean and variance** of $X \sim$ Gamma(α, β), $\alpha, \beta > 0$, β is the scale parameter.

   (a) **PDF**:

   $$f(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} x^{\alpha-1} e^{-\frac{x}{\beta}}$$

   $$x > 0$$

   $$\alpha > 0,\ \beta > 0$$

   (b) **MGF**: $M_X(t) = (1 - \beta t)^{-\alpha}$, $t < \frac{1}{\beta}$

   (c) **Mean and Variance**:

   $$E[X] = \alpha\beta,\ Var(X) = \alpha\beta^2$$

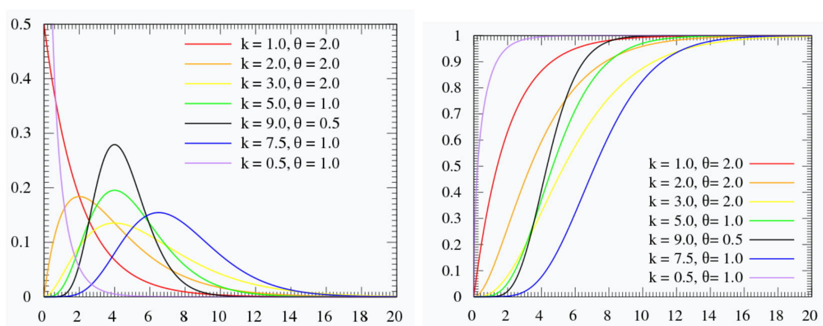

Figure 2.3: Gamma PDF (left) CDF (right)

2. **Additivity and Scaling**

   (a) If $X_i \overset{ind}{\sim}$ Gamma(α_i, β), then $\sum_{i=1}^k X_i \sim$ Gamma$(\sum_{i=1}^k \alpha_i, \beta)$

   (b) If $X \sim$ Gamma(α, β) then $cX \sim$ Gamma$(\alpha, c\beta)$

3. Random sample $X_1, ..., X_n \overset{i.i.d}{\sim}$ Gamma(α, β) where β is target parameter and α is known: exponential family? sufficient statistic? minimal sufficient statistic? complete statistic? Fisher information? UMVUE?

   (c) **Exponential family** form:

   $$\prod_{i=1}^n x_i^{\alpha-1} \cdot (\Gamma(\alpha)\beta^\alpha)^{-1} \exp\left(-\frac{1}{\beta}\sum_{i=1}^n x_i\right)$$

   (d) **Scale family**: the standard distribution is Gamma$(\alpha, 1)$ and $\frac{X_i}{\beta} \sim \Gamma(\alpha, 1)$

   (e) **Minimal sufficient and complete statistic**: $\sum_{i=1}^n X_i$

   (f) **Fisher information**: $\frac{n\alpha}{\beta^2}$

   (g) **UMVUE**: $\frac{1}{n\alpha}\sum_{i=1}^n X_i$

4. **Related Distributions**

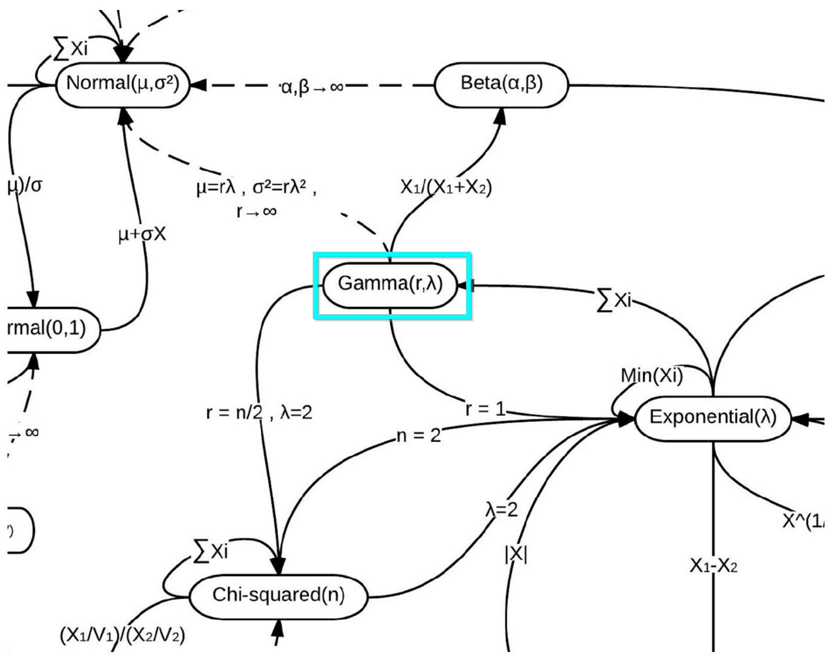

Figure 2.4: Gamma related distributions

(h) If $X \sim \text{Gamma}(\alpha, \beta)$, for $\alpha = 1$, $X \sim \text{Exp}(\beta)$

(i) **Pivot:** If $X \sim \text{Gamma}(\alpha, \beta)$, for $\alpha = \frac{n}{2}$, $\beta = 2$, $X \sim \chi_n^{2*}$

(j) If $X \sim \text{Gamma}(\alpha, \theta)$ and $Y \sim \text{Gamma}(\beta, \theta)$ are independent, then $Z = \frac{X}{X+Y} \sim \text{Beta}(\alpha, \beta)$

(k) If $X \sim \text{Gamma}(\alpha, \beta)$, for large α X converges in distribution to $\text{N}(\alpha\beta, \alpha\beta^2)$

(l) If $X \sim \text{Gamma}(\alpha, \beta)$, then $\frac{1}{X} \sim \text{InvGamma}(\alpha, \frac{1}{\beta})$

5. **Example problems** and key steps

(m) **Example** You are at a bank, and there are 3 people ahead of you. The serving time for each person is Exponential with mean 2 minutes. Only one person at a time can be served. The distribution of your waiting time until it's your turn to be served is $\text{Gamma}(3, 12)$.

6. Other notes

(n) Gamma is sometimes considered the continuous analog of the negative binomial distribution

2.2.1 Inverse Gamma

If $X \sim \text{Gamma}(\alpha, \beta)$, then $\frac{1}{X} \sim \text{InvGamma}(\alpha, \frac{1}{\beta})$

1. **PDF, CDF,MGF, mean and variance** of $X \sim \text{InvGamma}(\alpha, \beta)$, $\alpha, \beta > 0$, β is the scale parameter.

   (a) **PDF**:

$$f(x) = \frac{\beta^\alpha}{\Gamma(\alpha)}(x)^{-(\alpha+1)}e^{-\frac{\beta}{x}}$$

$$x > 0$$

$$\alpha > 1,\ \beta > 0$$

   (b) **CDF:**

$$\frac{\Gamma(\alpha, \frac{\beta}{x})}{\Gamma(\alpha)}$$

   (a) **MGF**: DNE

   (b) **Mean and Variance**:

$$E[X] = \frac{\beta}{\alpha - 1},\ Var(X) = \frac{\beta^2}{(\alpha-1)^2(\alpha-2)} \quad \alpha > 2$$

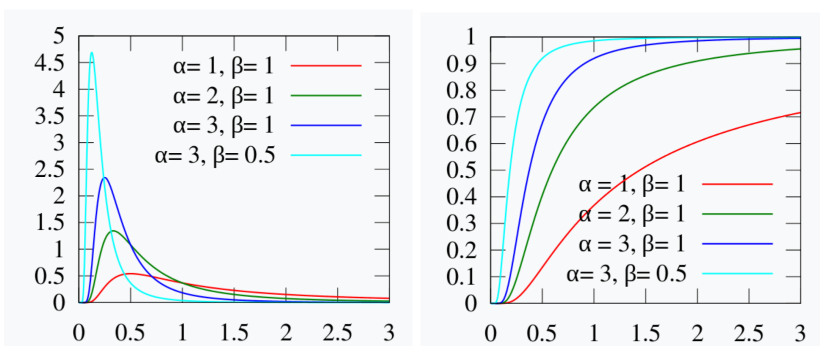

Figure 2.5: Inverse Gamma PDF (left) CDF (right)

2. **Scaling**

   (a) If $X \sim \text{InvGamma}(\alpha, \beta)$ then $cX \sim \text{Gamma}(\alpha, c\beta)$ for $c > 0$

3. Random sample $X_1, ..., X_n \overset{i.i.d}{\sim} \text{Gamma}(\alpha, \beta)$ where β is target parameter and α is known: exponential family? sufficient statistic? minimal sufficient statistic? complete statistic? Fisher information? UMVUE?

   (b) **Exponential family** form:

$$\prod_{i=1}^{n} x_i^{-(\alpha+1)} \cdot \left(\frac{\beta^\alpha}{\Gamma(\alpha)}\right) \exp\left(-\beta \sum_{i=1}^{n} \frac{1}{x_i}\right)$$

   (c) **Scale family**: the standard distribution is $\text{InvGamma}(\alpha, 1)$

   (d) **Minimal sufficient and complete statistic**: $\sum_{i=1}^{n} \frac{1}{X_i}$

   (e) **Fisher information**: TBD

(f) **UMVUE**: TBD

4. **Related Distributions**

(g) If $X \sim \text{Gamma}(\alpha, \beta)$, then $\frac{1}{X} \sim \text{InvGamma}(\alpha, \frac{1}{\beta})$

(h) If $X \sim \text{InvGamma}(1, c)$ then $\frac{1}{X} \sim EXP(c)$

2.3 Normal

1. **PDF, MGF, mean and variance** of $X \sim N(\mu, \sigma^2)$

   (a) **PDF**:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

$$-\infty < x < \infty$$

$$-\infty < \mu < \infty$$

$$\sigma > 0$$

   (b) **MGF**: $M_X(t) = \exp(\mu t + \frac{\sigma^2 t^2}{2})$

   (c) **Mean and Variance**:

$$E[X] = \mu,\ Var(X) = \sigma^2$$

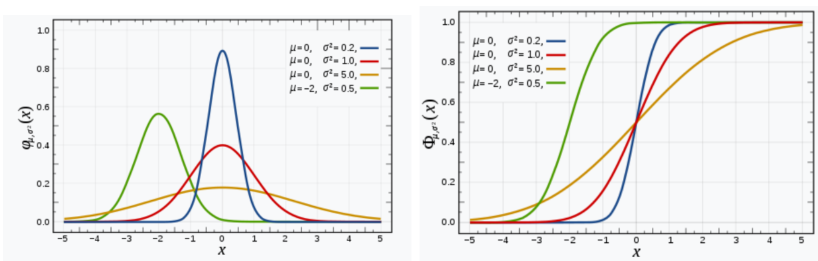

Figure 2.6: Normal PDF (left) CDF (right)

2. **Linearity** and **additivity**

   (a) If $X \sim N(\mu, \sigma^2)$, then $aX + b \sim N(a\mu + b, a^2\sigma^2)$

   (b) If $X_1 \sim N(\mu_1, \sigma_1^2)$ and $X_2 \sim N(\mu_2, \sigma_2^2)$ and X_1, X_2 are independent, then $X_1 \pm X_2 \sim N(\mu_1 \pm \mu_2, \sigma_1^2 + \sigma_2^2)$.

3. Population $X_1,...,X_n \overset{i.i.d}{\sim} N(\mu, \sigma_0^2)$ (μ is unknown and σ_0^2 is known)

   (c) **Exponential family** form: $(2\pi\sigma_0^2)^{-\frac{n}{2}} \exp\left(-n\frac{\mu^2}{2\sigma_0^2}\right) \exp\left(-\sum_{i=1}^n \frac{x_i^2}{2\sigma_0^2}\right) \exp\left(\frac{\mu}{\sigma_0^2}\sum_{i=1}^n x_i\right)$

   (d) **Location family**: the standard distribution is $N(0, \sigma_0^2)$ and $X_i - \mu \sim N(0, \sigma_0^2)$.

   (e) **Minimal sufficient and complete statistic**: $\sum_{i=1}^n X_i$

   (f) **Fisher information**: $\frac{n}{\sigma_0^2}$

   (g) **UMVUE**: $\frac{1}{n}\sum_{i=1}^n X_i$

   (h) **Conjugate prior** of $\mu \sim N(a, b^2)$

   $X|\mu \sim N(\mu, \sigma_0^2)$

   $\mu \sim N(a, b^2)$

   Posterior distribution $\mu|(X = x) \sim N\left(\left(\frac{a}{b^2} + \frac{n\bar{x}}{\sigma_0^2}\right)\left(\frac{b^2\sigma_0^2}{\sigma_0^2 + nb^2}\right), \left(\frac{b^2\sigma_0^2}{\sigma_0^2 + nb^2}\right)\right)$

4. Population $X_1,...,X_n \overset{i.i.d}{\sim} N(0, \sigma^2)$ (σ^2 is unknown)

   (i) **Exponential family** form:

$$(2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-\frac{1}{2\sigma^2}\sum_{i=1}^n x_i^2\right)$$

(j) **Scale family**: the standard distribution is $N(0,1)$ and $\frac{X_i}{\sigma} \sim N(0,1)$.

(k) **Minimal sufficient and complete statistic**: $\sum_{i=1}^n X_i^2$

(l) **Fisher information**: $\frac{n}{2\sigma^4}$

(m) **UMVUE**: $\frac{1}{n}\sum_{i=1}^n X_i^2$

(n) **Conjugate prior** of $\frac{1}{\sigma^2} \sim \text{Gamma}(\alpha, \beta)$

$X|\sigma^2 \sim N(0, \sigma^2)$

$\sigma^2 \sim \text{InvGamma}(\alpha, \beta)$

Posterior distribution $\sigma^2|(X = x) \sim \text{InvGamma}\left(\alpha + \frac{n}{2}, \frac{1}{\frac{1}{2}\sum_{i=1}^n X_i^2 + \frac{1}{\beta}}\right)$

5. Population $X_1,...,X_n \overset{i.i.d}{\sim} N(\mu, \sigma^2)$ (μ is unknown and σ^2 is unknown)

(o) **Exponential family** form:

$$(2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-n\frac{\mu^2}{2\sigma^2}\right) \exp\left(\frac{\mu}{\sigma^2}\sum_{i=1}^n x_i - \frac{1}{2\sigma^2}\sum_{i=1}^n x_i^2\right)$$

(p) **Location-Scale family**: the standard distribution is $N(0,1)$ and $\frac{X_i - \mu}{\sigma} \sim N(0,1)$.

(q) **Minimal sufficient and complete statistic**: $\left(\sum_{i=1}^n X_i, \sum_{i=1}^n X_i^2\right)$

(r) **Fisher information**: $I = \text{diag}(\frac{n}{\sigma^2}, \frac{n}{2\sigma^4})$

(s) **UMVUE**: $\left(\bar{X}, S^2\right)$

(t) **Conjugate prior** of $\mu \sim N(a, b^2)$, $\sigma^2 \sim \text{InvGamma}(\alpha, \beta)$ or $\frac{1}{\sigma^2} \sim \text{Gamma}(\alpha, \beta)$

6. **Related Distributions**

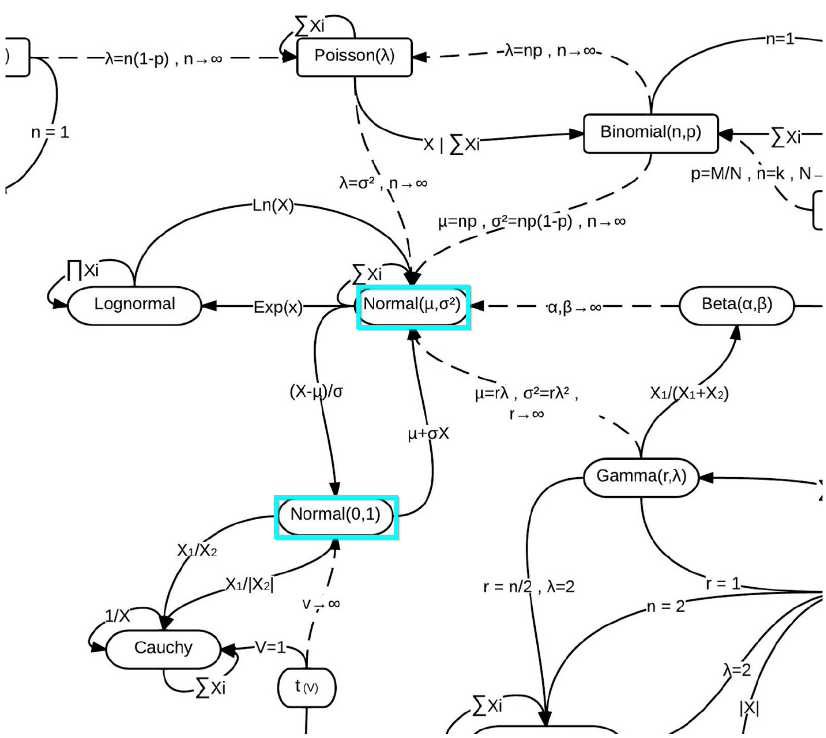

Figure 2.7: Normal related distributions

(u) For $X_1, ..., X_n \overset{i.i.d}{\sim} N(0,1)$, $\sum_{i=1}^n X_i^2 \sim \chi_n^2$

(v) For $X_1, ..., X_n \overset{i.i.d}{\sim} N(\mu, \sigma^2)$:

i. $\bar{X} \sim N(\mu, \frac{\sigma^2}{n})$

ii. $\frac{S^2}{\sigma^2} \sim \chi_{n-1}^2$ where $S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2$

iii. $\bar{X}$ and S^2 are independent and $\frac{\bar{X}-\mu}{\sqrt{S^2/n}} \sim t_{n-1}$

(w) If $Z \sim N(0,1)$, $Y \sim \chi_n^2$ and they are independent, then $X = \frac{Z}{\sqrt{Y/n}} \sim t_n$

(x) If $Y \sim \chi_n^2$, $Z \sim \chi_m^2$ and they are independent, then $X = \frac{Y}{Z} \sim F_{n,m}$

(y) If $X \sim N(\mu, \sigma^2)$, then $\exp(X) \sim \log N(\mu, \sigma^2)$

(z) If $X_1 \sim N(0,1)$ and $X_2 \sim N(0,1)$ are independent then $\frac{X_1}{X_2} \sim \text{Cauchy}(0,1)$

7. **Example problems** and key steps

2.4 Exponential

Let us say that X is distributed EXP(β). We know the following: You're sitting on an open meadow right before the break of dawn, wishing that airplanes in the night sky were shooting stars, because you could really use a wish right now. You know that shooting stars come on average every 15 minutes, but a shooting star is not "due" to come just because you've waited so long. Your waiting time is memoryless; the additional time until the next shooting star comes does not depend on how long you've waited already.

1. **PDF, CDF, MGF, mean and variance** of $X \sim \text{Exp}(\beta)$, $\beta > 0$

   (a) **PDF**:

   $$f(x) = \frac{1}{\beta}e^{-\frac{x}{\beta}}$$

   $$x > 0$$

   (b) **CDF**: $f(x) = (1 - e^{-\frac{x}{\beta}}) \cdot I_{[0,\infty)}(x)$

   (c) **MGF**: $M_X(t) = (1 - \beta t)^{-1}$

   (d) **Mean and Variance**:

   $$E[X] = \beta, \; Var(X) = \beta^2$$

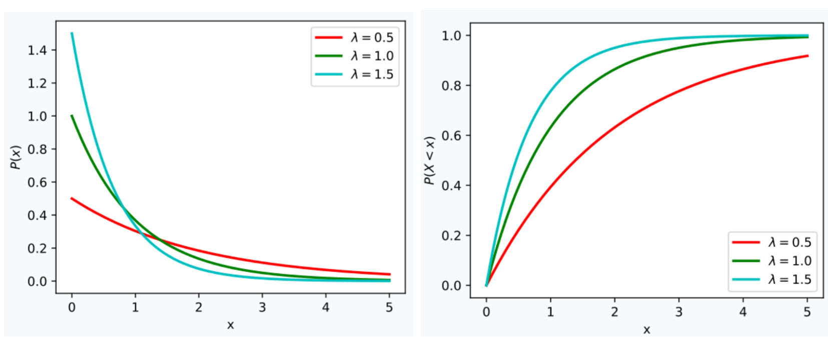

Figure 2.8: Exponential PDF (left) CDF (right)

2. Random sample $X_1, ..., X_n \sim \text{Exp}(\beta)$ where β is target parameter: exponential family? sufficient statistic? minimal sufficient statistic? complete statistic? Fisher information? UMVUE?

   (a) **Exponential family**:

   $$\left(\frac{1}{\beta}\right)^n \exp\left(-\frac{1}{\beta}\sum_{i=1}^{n} x_i\right)$$

   (b) **Scale family**: The standard distribution is Exp(1) and $\frac{X_i}{\beta} \sim \text{Exp}(1)$

   (c) **Minimal sufficient and complete statistic**: $\sum_{i=1}^{n} X_i$

   (d) **Fisher information**: $\frac{n}{\beta^2}$

   (e) **UMVUE**: $\frac{1}{n}\sum_{i=1}^{n} X_i$

3. **Related Distributions**

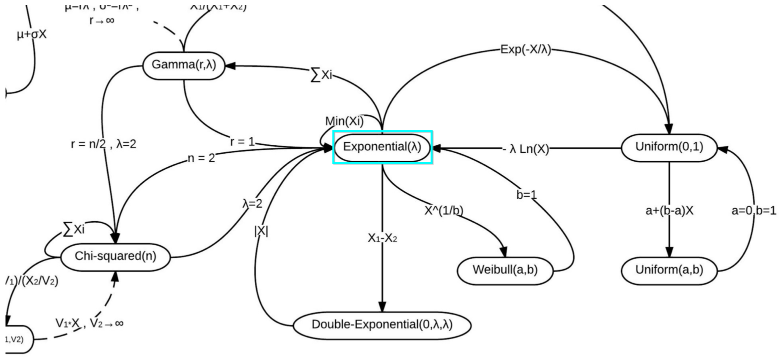

Figure 2.9: Exponential and related distributions

- If $X \sim \text{Laplace}(\mu, \frac{1}{\beta})$, then $|X - \mu| \sim \text{Exp}(\beta)$ (Laplace is double exponential)
- $\lambda_1 X_1 - \lambda_2 X_2 \sim \text{Laplace}(0,1)$*
- If $X \sim \text{Pareto}(1, \lambda)$, then $\log(X) \sim \text{Exp}(\lambda)$
- If $X_i \sim \text{Unif}(0,1)$, then $\lim\limits_{n\to\infty} n \min(X_1, ..., X_n) \sim \text{Exp}(1)$*
- Limit of a scaled Beta distribution $\lim\limits_{n\to\infty} n\text{Beta}(1,n) \sim \text{Exp}(1)$

If $X \sim \text{Exp}(\lambda)$ and $X_i \sim \text{Exp}(\lambda_i)$

- $kX \sim \text{Exp}\left(\frac{\lambda}{k}\right)$, closure under scaling by a positive factor*
- $ke^x \sim \text{Pareto}\,(k, \lambda)$
- $e^{-X} \sim \text{Beta}\,(\lambda, 1)$
- $\sqrt{X} \sim \text{Rayleigh}\left(\frac{1}{\sqrt{2\lambda}}\right)$
- $X \sim \text{Weibull}\left(\frac{1}{\lambda}, 1\right)$
- $X^2 \sim \text{Weibull}\left(\frac{1}{\lambda^2}, \frac{1}{2}\right)$
- $\min(X_1, ..., X_n) \sim \text{Exp}(\lambda_1, ..., \lambda_n)$

If also $\lambda_i = \lambda$

- $\sum_{i=1}^{k} X_i \sim \text{Gamma}(k, \lambda)$ or if β parameter then $\sim \text{Gamma}(k, \frac{1}{\beta})$*
  - $X \sim \text{Exp}(\beta)$ is equivalent to $X \sim \text{Gamma}(1, \beta)$
- T=$\sum_{i=1}^{n} X_i$, then $2\lambda T \sim \chi^2_{2n}$*
- $X_i - X_j \sim \text{Laplace}(0, \lambda^{-1})$

If also X_i are independent, then

- $\frac{X_i}{X_i + X_j} \sim \text{Unif}(0,1)$*
- $Z = \frac{\lambda_i X_i}{\lambda_j X_j}$has *pdf* $f_Z(z) = \frac{1}{(z+1)^2}$. This can be used as a confidence interval for $\frac{\lambda_i}{\lambda_j}$

If also $\lambda = \frac{1}{2}$ then $X \sim \chi^2_2$, a chi squared distribution with two degrees of freedom, hence*

- $\text{Exp}(\lambda) = \frac{1}{2\lambda}\text{Exp}(\frac{1}{2}) \sim \frac{1}{2\lambda}\chi^2_2$ so $\sum_{i=1}^{n} \text{Exp}(\lambda) \sim \frac{1}{2\lambda}\chi^2_2$*
- If $X \sim \text{Exp}(\frac{1}{\lambda})$ and $Y|X \sim \text{Pois}(X)$ then $Y \sim \text{Geom}(\frac{1}{1+\lambda})$

4. **Example problems** and key steps

(f) **Example** The waiting time until the next shooting star is distributed EXP(4) hours. Here $\beta = 4$ is the rate parameter, since shooting stars arrive at a rate of 1 per 1/4 hour on average. The expected time until the next shooting star is $1/\beta = 1/4$ hour.

(g) **Memorylessness** The Exponential Distribution is the only continuous memoryless distribution. The memoryless property says that for $X \sim \text{EXP}(\lambda)$and any positive numbers s and t,

$$P(X > s + t | X > s) = P(X > t)$$

Equivalently,

$$X - a | (X > a) \sim \text{EXP}(\lambda)$$

For example, a product with an EXP(λ) lifetime is always "as good as new" (it doesn't experience wear and tear). Given that the product has survived a years, the additional time that it will last is still EXP(λ).

2.5 Beta

1. **PDF, MGF, mean and variance** of $X \sim \text{Beta}(\alpha, \beta)$, $\alpha, \beta > 0$, α, β are shape parameters.

   (a) **PDF**:

$$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{\mathrm{B}(\alpha,\beta)}$$

$$1 > x > 0$$

$$\alpha > 0,\ \beta > 0$$

   (b) **Mean and Variance**:

$$E[X] = \frac{\alpha}{\alpha+\beta},\ Var(X) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$$

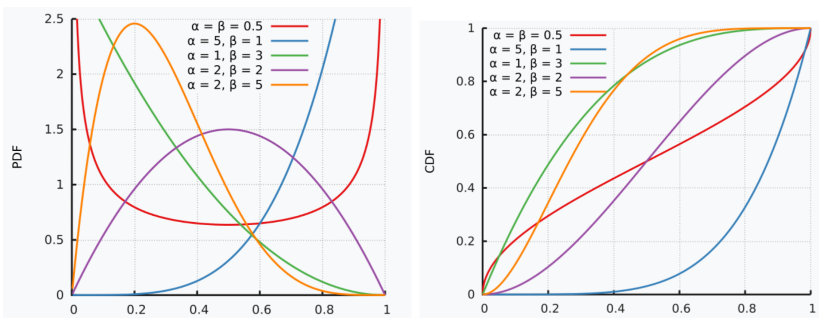

Figure 2.10: Beta PDF (left) CDF (right)

2. Random sample $X_1, ..., X_n \overset{i.i.d}{\sim} \text{Beta}(\alpha, \beta)$ where α, β are target parameters: exponential family? sufficient statistic? minimal sufficient statistic? complete statistic? Fisher information? UMVUE?

   (a) **Exponential family** form:

$$\exp\left[(\alpha-1)\sum_{i=1}^{n}\log x_i + (\beta-1)\sum_{i=1}^{n}\log(1-x_i) + n\log\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\right]$$

   (b) **Minimal sufficient and complete statistic**: $T = (\sum_{i=1}^{n}\log X_i, \sum_{i=1}^{n}\log(1-X_i))$

   (c) **Conjugate prior** to the Binomial
   $X|p \sim \text{Bin}(n, p)$
   $p \sim \text{Beta}(a, b)$
   Posterior distribution: $p|(X = x) \sim \text{Beta}(a + x, b + n - x)$

   (d) For β known, the **MLE** of α $\sum_{i=1}^{n}\log X_i$

   (e) For α known, the **MLE** of β $\boldsymbol{TBD} ****$

3. **Related Distributions**

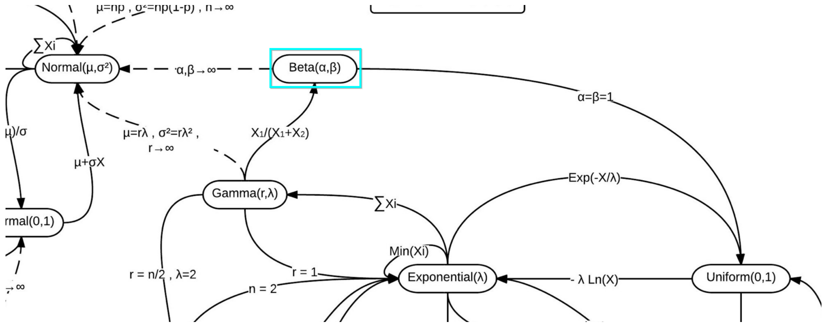

Figure 2.11: Beta related distributions

(f) If $X \sim \text{Beta}(\alpha,\beta)$ then $1-X \sim \text{Beta}(\beta,\alpha)$

(g) **Pivot:** If $X \sim \text{Beta}(\frac{n}{2},\frac{m}{2})$ with $n>0, m>0$, then $\frac{mX}{n(1-X)} \sim F(n,m)$

(h) If $X \sim \text{Beta}(\alpha,1)$ then $-\ln(X) \sim \text{EXP}(\alpha)$

(i) For large n, if $X_n \sim \text{Beta}(n\alpha,n\beta)$ then $\sqrt{n}(\bar{X}_n - \frac{a}{a+\beta}) \xrightarrow{d} N(0,\frac{a\beta}{(a+\beta)^3})$, or similarly $\text{Beta}(n\alpha,n\beta) \xrightarrow{d} N(\frac{a}{a+\beta},\frac{a\beta}{(a+\beta)^3})$

(j) If $X \sim \chi^2_\alpha$ and $Y \sim \chi^2_\beta$ are independent, then $\frac{X}{X+Y} \sim \text{Beta}(\frac{\alpha}{2},\frac{\beta}{2})$

(k) If $X \sim \text{Unif}(0,1)$ and $\alpha>0$ then $X^{\frac{1}{\alpha}} \sim \text{Beta}(\alpha,1)$

(l) If $X \sim \text{Unif}(a,b)$ then single order statistics (Casella example 5.4.5): $\frac{X_{(j)}-a}{b-a} \sim \text{beta}(j,n-j+1)$*

(m) If $X \sim \text{Gamma}(\alpha,\theta)$ and $Y \sim \text{Gamma}(\beta,\theta)$ are independent, then $Z = \frac{X}{X+Y} \sim \text{Beta}(\alpha,\beta)$*

(n) If $X \sim \text{Cauchy}(0,1)$ then $\frac{1}{1+X^2} \sim \text{Beta}(\frac{1}{2},\frac{1}{2})$

4. **Example problems** and key steps

- Determine a minimal sufficient statistic if $\alpha = 2\beta$
  $\exp\left[(2\beta-1)\sum_{i=1}^n \log x_i + (\beta-1)\sum_{i=1}^n \log(1-x_i) + n\log\frac{\Gamma(3\beta)}{\Gamma(2\beta)\Gamma(\beta)}\right]$
  $= \exp\left[\sum_{i=1}^n \log x_i + (\beta-1)\sum_{i=1}^n \log x_i^2(1-x_i) + n\log\frac{\Gamma(3\beta)}{\Gamma(2\beta)\Gamma(\beta)}\right]$,
  which is exponential family with minimal sufficient statistic $\sum_{i=1}^n \log x_i^2(1-x_i) = 2T_1 + T_2$ where $T = (\sum_{i=1}^n \log X_i, \sum_{i=1}^n \log(1-X_i))$
- Determine a minimal sufficient statistic if $\alpha = \beta^2$
  $p_\beta(x) = \exp\left[(\beta^2-1)T_1(x) + (\beta-1)T_2(x) + n\log\frac{\Gamma(\beta+\beta^2)}{\Gamma(\beta)\Gamma(\beta^2)}\right]$ use Lehmann Scheffe to show min sufficient that $T(x) = T(y)$

5. **Other notes**

2.6 Log-normal*

In probability theory, a log-normal (or lognormal) distribution is a continuous probability distribution of a random variable whose logarithm is normally distributed. Thus, if the random variable X is log-normally distributed, then $Y = \ln(X)$ has a normal distribution. Equivalently, if Y has a normal distribution, then the exponential function of Y, $X = \exp(Y)$, has a log-normal distribution. A random variable which is log-normally distributed takes only positive real values. It is a convenient and useful model for measurements in exact and engineering sciences, as well as medicine, economics and other topics (e.g., energies, concentrations, lengths, prices of financial instruments, and other metrics).

1. **PDF, MGF, mean and variance** of $X \sim \mathcal{L}N(\mu, \sigma^2)$
   (a) **PDF**:
   $$f(x) = \frac{1}{x\sqrt{2\pi\sigma^2}} e^{-\frac{(\log x - \mu)^2}{2\sigma^2}}$$
   $$0 < x < \infty$$
   $$-\infty < \mu < \infty$$
   $$\sigma > 0$$
   (b) **MGF**: DNE
   (c) **Mean and Variance**:
   $$E[X] = \theta = e^{\mu + \frac{\sigma^2}{2}},\ Var(X) = \theta^2(e^{\sigma^2} - 1)$$

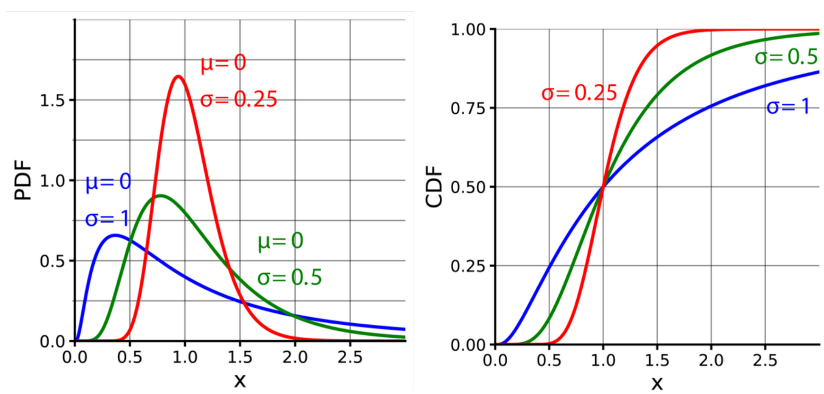

Figure 2.12: Log Normal PDF (left) CDF (right)

2. **Multiple, Reciprocal, Power**
   (a) If $X \sim \mathcal{L}N(\mu, \sigma^2)$, then $aX \sim \mathcal{L}N(\mu + \ln a, \sigma^2)$ for $a > 0$
      i. If $X \sim \mathcal{L}N(\mu, \sigma^2)$, then $\frac{1}{X} \sim \mathcal{L}N(-\mu, \sigma^2)$
      ii. If $X \sim \mathcal{L}N(\mu, \sigma^2)$, then If $X^a \sim \mathcal{L}N(a\mu, a^2\sigma^2)$
      iii. If $X_1, ..., X_n \overset{i.i.d}{\sim} \mathcal{L}N(\mu_i, \sigma_i^2)$ then $Y = \prod\limits_{i=1}^{n} X_i \sim \mathcal{L}N(\sum\limits_{i=1}^{n} \mu_i, \sum\limits_{i=1}^{n} \sigma_i^2)$

3. Population X_1,...,$X_n \overset{i.i.d}{\sim} \mathcal{L}N(\mu, \sigma^2)$ (μ is unknown and σ^2 is unknown)
   (a) **Exponential family** form: $(x)^{-\frac{n}{2}}(2\pi\sigma^2)^{-\frac{n}{2}} \exp\left(-n\frac{\mu^2}{2\sigma^2}\right) \exp\left(\frac{\mu}{\sigma^2}\sum_{i=1}^{n} \log x_i - \frac{1}{2\sigma^2}\sum_{i=1}^{n} \log x_i^2\right)$
   (b) The standard distribution is $\mathcal{L}N(0, 1)$

(c) **Minimal sufficient and complete statistic**: $\left(\sum_{i=1}^n \log X_i, \sum_{i=1}^n \log X_i^2\right)$

(d) **Fisher information**: $I = \text{diag}(\frac{1}{\sigma^2}, \frac{2}{\sigma^2})$

(e) **MLE** $\hat{\mu} = \dfrac{\sum_{i=1}^n \log X_i}{n}$, $\hat{\sigma}^2 = \dfrac{\sum_{i=1}^n (\log X_i - \hat{\mu})^2}{n}$

2.7 Weibull*

In probability theory and statistics, the Weibull distribution is a continuous probability distribution. It models a broad range of random variables, largely in the nature of a time to failure or time between events. Examples are maximum one-day rainfalls and the time a user spends on a web page.

1. **PDF, CDF, MGF, mean and variance** of $X \sim \text{Weibull}(\lambda, k)$,

   (a) **PDF**:
   $$f(x) = \frac{k}{\lambda}\left(\frac{x}{\lambda}\right)^{k-1} e^{-(\frac{x}{\lambda})^k}$$
   $$x \geq 0,\ \lambda > 0 \text{ is scale parameter, } k > 0 \text{ is shape parameter}$$

   (b) **CDF**: $f(x) = (1 - e^{-(\frac{x}{\lambda})^k}) \cdot I_{[0,\infty)}(x)$

   (c) **MGF**: $M_X(t) = \sum_{n=0}^{\infty} \frac{t^n \lambda^n}{n!}\Gamma(1 + \frac{n}{k}),\ k \geq 1$

   (d) **Mean and Variance**:
   $$E[X] = \lambda\Gamma(1 + \frac{1}{k}),\ Var(X) = \lambda^2\left[\Gamma(1 + \frac{2}{k}) - \left(\Gamma(1 + \frac{1}{k})\right)^2\right]$$

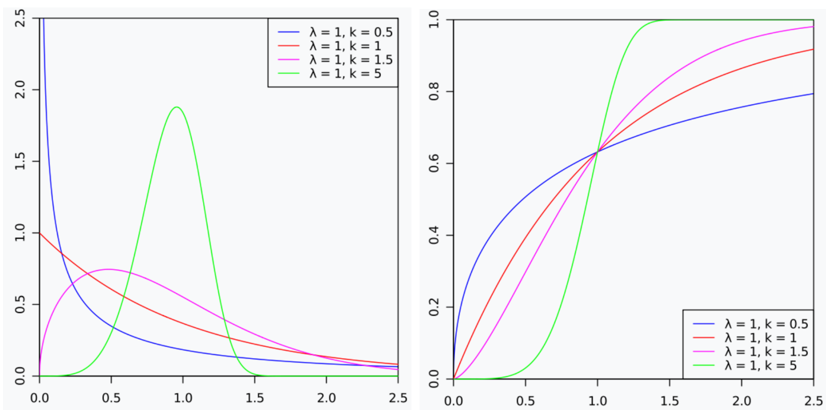

Figure 2.13: Weibull PDF (left) CDF (right)

2. Random sample $X_1, ..., X_n \sim \text{Weibull}(\lambda, k)$ where λ is target parameter: exponential family? sufficient statistic? minimal sufficient statistic? complete statistic? Fisher information? UMVUE?

   k unknown

   Not exponential family for λ, k unknown

   $\log f(x) = \log(\frac{k}{\lambda}) + (k-1)\log(\frac{x}{\lambda}) - (\frac{x}{\lambda})^k$

   If k known

   (a) **Exponential family** form: $f(x) = \frac{k}{\lambda^k}(x)^{k-1} e^{-(\frac{1}{\lambda} * x)^k}$

   (b) **Scale family**: The standard distribution is Exp(1) and $\frac{X_i}{\beta} \sim \text{Exp}(1)$

   (c) **Minimal sufficient and complete statistic**: $\sum_{i=1}^{n} X_i^k$

   (d) **Fisher information**: $\frac{nk^2}{\lambda^2}$

3. **Related Distributions**

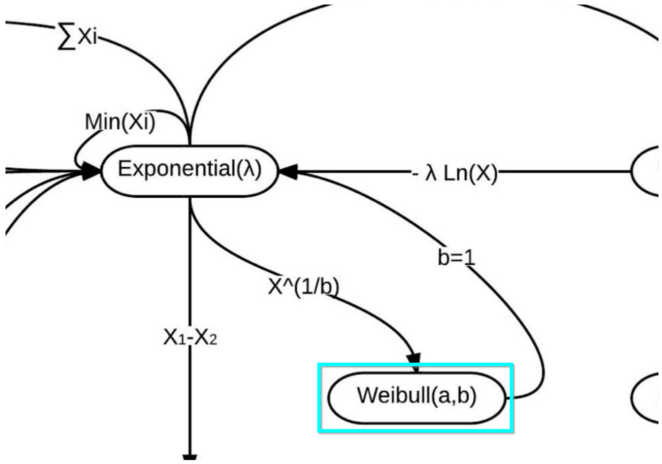

Figure 2.14: Weibull and related distributions

- If $X \sim \text{Weibull}(\lambda, \frac{1}{2})$ then $\sqrt{X} \sim \text{EXP}(\frac{1}{\sqrt{\lambda}})$

2.8 Dirichlet**

In probability and statistics, the Dirichlet distribution (after Peter Gustav Lejeune Dirichlet), often denoted $\mathrm{Dir}(\boldsymbol{\alpha})$, is a family of continuous multivariate probability distributions parameterized by a vector $\boldsymbol{\alpha}$ of positive reals. It is a multivariate generalization of the beta distribution, hence its alternative name of multivariate beta distribution (MBD). Dirichlet distributions are commonly used as prior distributions in Bayesian statistics, and in fact, the Dirichlet distribution is the conjugate prior of the categorical distribution and multinomial distribution.

1. **PDF, MGF, mean and variance** of $(X_1, ..., X_K) \sim \mathrm{Dir}(\alpha_1, ..., \alpha_k)$

   (a) **PDF**:

   $$\frac{1}{\mathrm{B}(\boldsymbol{\alpha})} \prod_{i=1}^{K} x_i^{\alpha_i - 1}$$

   $$\sum_{i=1}^{K} x_i = 1$$

   $$\mathrm{B}(\boldsymbol{\alpha}) = \frac{\prod_{i=1}^{K} \Gamma(\alpha_i)}{\Gamma(\alpha_0)} \text{ where } \alpha_0 = \sum_{i=1}^{K} \alpha_i$$

   (b) **Mean and Variance**:

   $$E[X_i] = \frac{\alpha_i}{\alpha_0},\ Var(X_i) = \frac{\tilde{\alpha_i}(1-\tilde{\alpha_i})}{\alpha_0+1} \text{ where } \tilde{\alpha_i} = \frac{\alpha_i}{\alpha_0}$$

   (a) **Covariance**: $\frac{-\alpha_i \alpha_j}{\alpha_0(\alpha_0+1)}$ for $i \neq j$

2. **Conjugate prior** exists but not going to be useful here

3. **Related Distributions**

   (a) For K independent r.v.s $Y_1 \sim \mathrm{Gamma}(\alpha_1, \beta)$,..., $Y_K \sim \mathrm{Gamma}(\alpha_K, \beta)$, we have $V = \sum_{i=1}^{K} Y_i \sim \mathrm{Gamma}(\alpha_0, \beta)$, $X = (X_1, ..., X_k) = \left(\frac{Y_1}{V}, ..., \frac{Y_K}{V}\right) \sim \mathrm{Dir}(\alpha_1, ..., \alpha_K)$

4. **Example problems** and key steps

5. Other notes

   (b) Multivariate generalization of Beta distribution

2.9 Inverse Gaussian*****

The inverse Gaussian distribution has several properties analogous to a Gaussian distribution. The name can be misleading: it is an "inverse" only in that, while the Gaussian describes a Brownian motion's level at a fixed time, the inverse Gaussian describes the distribution of the time a Brownian motion with positive drift takes to reach a fixed positive level.

Its cumulant generating function (logarithm of the characteristic function) is the inverse of the cumulant generating function of a Gaussian random variable.

1. **PDF, MGF, mean and variance** of $X \sim \mathrm{IG}(\mu, \lambda)$

   (a) **PDF**:

   $$f(x) = \sqrt{\frac{\lambda}{2\pi x^3}} \exp\left[-\frac{\lambda(x-\mu)^2}{2\mu^2 x}\right]$$

   $$0 < x < \infty$$

   $$\mu > 0$$

   $$\lambda > 0$$

   (b) **MGF**: $M_X(t) = \exp\left[\frac{\lambda}{\mu}\left(1 - \sqrt{1 - \frac{2\mu^2 t}{\lambda}}\right)\right]$

   (c) **Mean and Variance**:

   $$E[X] = \mu,\ Var(X) = \frac{\mu^3}{\lambda}$$

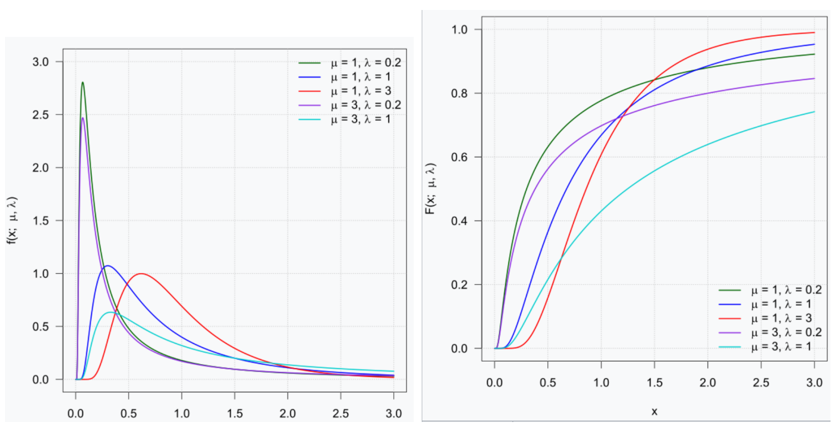

Figure 2.15: Inverse Gaussian PDF (left) CDF (right)

2. **Sum and Scaling**

   (a) If $X_i \sim \mathrm{IG}(\mu_0 w_i, \lambda_0 w_i^2)$ then $S = \sum_{i=1}^{n} X_i \sim \mathrm{IG}(\mu_0 \sum_{i=1}^{n} w_i, \lambda_0 (\sum_{i=1}^{n} w_i)^2)$, under condition that $\frac{\mathrm{Var}(X_i)}{\mathrm{E}(X_i)} = \frac{\mu_0^2 w_i^2}{\lambda_0 w_i^2} = \frac{\mu_0^2}{\lambda_0}$ is constant for all i

   (b) If $X \sim \mathrm{IG}(\mu, \lambda)$, $tX \sim \mathrm{IG}(t\mu, t\lambda)$ for $t > 0$

3. Population $X_1, \ldots, X_n \overset{i.i.d}{\sim} \mathrm{IG}(\mu, \lambda)$

   (a) **Exponential family** form: $f(x) = \exp\left[-\frac{\lambda}{2\mu^2}x + \frac{\lambda}{\mu} - \frac{\lambda}{2x} + \frac{1}{2}\log\lambda + \frac{1}{2}\log 2\pi x^3\right]$

   (b) **Minimal sufficient and complete statistic**: $\left(\sum_{i=1}^{n} X_i, \sum_{i=1}^{n} \frac{1}{X_i}\right)$

(c) **Fisher information**: $I_n(\mu, \lambda) = \left(\frac{n\lambda}{\mu^3}, \frac{n}{2\lambda^2}\right)$

(d) **MLE** $\hat{\mu} = \bar{X}$, $\hat{\lambda} = \dfrac{n}{\sum\limits_{i=1}^{n} \left(\frac{1}{X_i} - \frac{1}{\bar{X}}\right)}$

4. **Related Distributions**

(e) **Pivot:** If $X \sim \mathrm{IG}(\mu, \lambda)$, then $\dfrac{\lambda(X-\mu)^2}{\mu^2 X} \sim \chi_1^2$

(f) If $X_i \sim \mathrm{IG}(\mu, \lambda)$ then $\sum\limits_{i=1}^{n} X_i \sim \mathrm{IG}(n\mu, n^2\lambda)$

(g) If $X_i \sim \mathrm{IG}(\mu, \lambda)$ then $\bar{X} \sim \mathrm{IG}(\mu, n\lambda)$

(h) If $X_i \sim \mathrm{IG}(\mu_i, 2\mu_i^2)$ then $\sum\limits_{i=1}^{n} X_i \sim \mathrm{IG}\left(\sum\limits_{i=1}^{n} \mu_i, 2\left(\sum\limits_{i=1}^{n} \mu_i\right)^2\right)$

2.10 Cauchy*

Cauchy is the distribution of the ratio of two independent normally distributed random variables with mean zero.

The Cauchy distribution is often used in statistics as the canonical example of a "pathological" distribution since both its expected value and its variance are undefined (but see § Moments below). The Cauchy distribution does not have finite moments of order greater than or equal to one; only fractional absolute moments exist.[1] The Cauchy distribution has no moment generating function.

1. **PDF, MGF, mean and variance** of $X \sim \text{Cauchy}(x_0, \gamma)$; x_0 location parameter, γ scale parameter

   (a) **PDF**:

   $$f(x) = \frac{1}{\pi\gamma\left[1 + \left(\frac{x-x_0}{\gamma}\right)^2\right]}$$

   $$-\infty < x < \infty$$

   $$-\infty < x_0 < \infty$$

   $$\gamma > 0$$

   (b) **CDF**: $F(x) = \frac{1}{\pi}\arctan\left(\frac{x-x_0}{\gamma}\right) + \frac{1}{2}$

   (c) **MGF**: Does not exist

   (d) **Mean and Variance**:

   $$E[X] = \text{undefined}, \; Var(X) = \text{undefined}$$

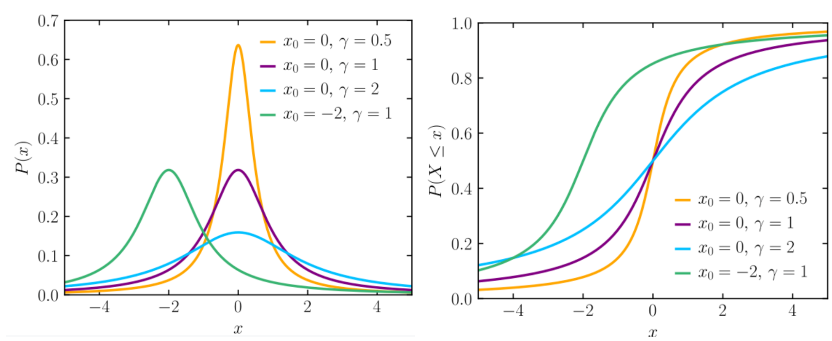

Figure 2.16: Cauchy PDF (left) CDF (right)

2. **Linearity** and **additivity**

   (a) If $X \sim \text{Cauchy}(x_0, \gamma)$, then $kX + \ell \sim \text{Cauchy}(x_0 k + \ell, \gamma|k|)$

   (b) If $X \sim \text{Cauchy}(x_0, \gamma_0)$ and $Y \sim \text{Cauchy}(x_1, \gamma_1)$ are independent, then $X + Y \sim \text{Cauchy}(x_0 + x_1, \gamma_0 + \gamma_1)$ and $X - Y \sim \text{Cauchy}(x_0 - x_1, \gamma_0 + \gamma_1)$

   (c) If $X \sim \text{Cauchy}(0, \gamma)$, then $\frac{1}{X} \sim \text{Cauchy}(0, \frac{1}{\gamma})$

   (d) If $X_1, ..., X_n \overset{i.i.d.}{\sim} \text{Cauchy}(0, 1)$ standard Cauchy distributed, then the sample mean $\frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X} \sim \text{Cauchy}(0, 1)$ is also standard Cauchy.

3. Population $X_1, ..., X_n \sim \text{Cauchy}(x_0, \gamma)$ (μ is unknown and σ_0^2 is known)*****

4. **Related Distributions**

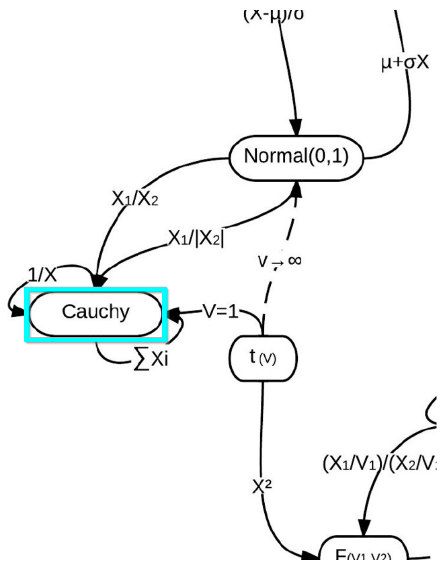

Figure 2.17: Cauchy related distributions

(e) If $X_1 \sim N(0,1)$ and $X_2 \sim N(0,1)$ are independent then $\frac{X_1}{X_2} \sim \text{Cauchy}(0,1)$

(f) $\text{Cauchy}(0,1) \sim t(\text{df} = 1)$ Student's t distribution

(g) $\text{Cauchy}(\mu, \sigma) \sim t_{(\text{df}=1)}(\mu, \sigma)$

(h) If $X \sim \text{Unif}(0,1)$ then $\tan(\pi(X - \frac{1}{2})) \sim \text{Cauchy}(0,1)$

(i) If $X \sim \text{Cauchy}(x_0, \gamma)$, then $\frac{1}{X} \sim \text{Cauchy}(\frac{x_0}{x_0^2+\gamma^2}, \frac{\gamma}{x_0^2+\gamma^2})$

5. **Example problems** and key steps

2.11 Laplace (Double Exponential)

In probability theory and statistics, the Laplace distribution is a continuous probability distribution named after Pierre-Simon Laplace. It is also sometimes called the double exponential distribution, because it can be thought of as two exponential distributions (with an additional location parameter) spliced together along the abscissa, although the term is also sometimes used to refer to the Gumbel distribution. The difference between two independent identically distributed exponential random variables is governed by a Laplace distribution, as is a Brownian motion evaluated at an exponentially distributed random time.

1. **PDF, CDF, MGF, mean and variance** of $X \sim \text{Laplace}(\mu, b)$; μ is location parameter, b is scale parameter.

   (a) **PDF**:
   $$f(x) = \frac{1}{2b} \exp\left(-\frac{|x-\mu|}{b}\right)$$
   $$-\infty < x < \infty,\ -\infty < \mu < \infty\ ,\ b > 0$$

   (b) **CDF**: $f(x) = \begin{cases} \frac{1}{2}\exp\left(\frac{x-\mu}{b}\right) & x \leq \mu \\ 1 - \frac{1}{2}\exp\left(-\frac{x-\mu}{b}\right) & x > \mu \end{cases}$

   (c) **MGF**: $M_X(t) = \frac{\exp(\mu t)}{1-b^2t^2}$, $|t| < \frac{1}{b}$

   (d) **Mean and Variance**:
   $$E[X] = \mu,\ Var(X) = 2b^2$$

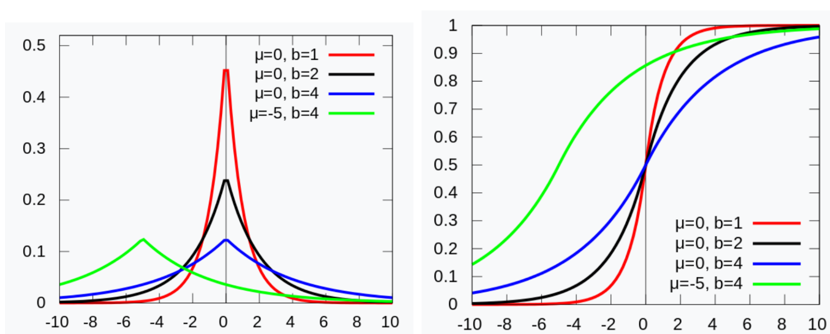

Figure 2.18: Laplace PDF (left) CDF (right)

2. Random sample $X_1, ..., X_n \sim \text{Laplace}(\mu, b)$ where μ, b is target parameter: exponential family? sufficient statistic? minimal sufficient statistic? complete statistic? Fisher information? UMVUE?

   (a) **Not exponential if** b **or** μ **unknown. If** μ **known exponential family** form: $f(x) = \frac{1}{2b}\exp\left(-\frac{1}{b}\sum_{i=1}^{n}|x-\mu|\right)$

   (b) **Location-Scale family**: The standard distribution is Laplace$(0,1)$ and $\frac{X-\mu}{b} \sim \text{Laplace}(0,1)$

   (c) **Minimal sufficient and complete statistic**: $\sum_{i=1}^{n}|X_i - \mu|$

   (d) **Fisher information**: for $b = 1$ $I_n(\mu) = n$, for $\mu = 0$ $I_n(b) = \frac{n}{b^2}$

3. **Related Distributions**

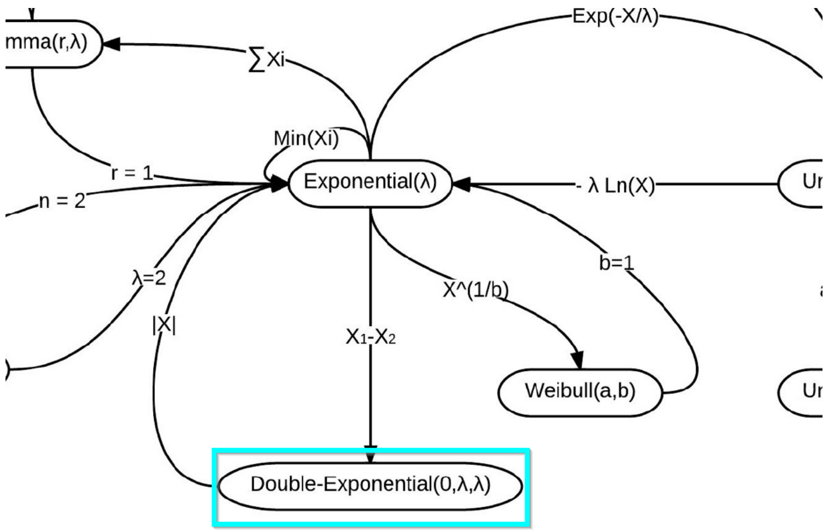

Figure 2.19: Laplace and related distributions

- If $X \sim \text{Laplace}(\mu, b)$ then $kX + c \sim \text{Laplace}(k\mu + c, |k|b)$
- If $X \sim \text{Laplace}(0, 1)$ then $bX \sim \text{Laplace}(0, b)$
- If $X \sim \text{Laplace}(0, b)$ then $|X| \sim \text{EXP}(b^{-1})$
- If $X \sim \text{Laplace}(\mu, b)$ then $|X - \mu| \sim \text{EXP}(b^{-1})$
- If $X, Y \sim \text{EXP}(\lambda)$, then $X - Y \sim \text{Laplace}(0, \lambda^{-1})$
- If $X_1, .., X_4 \sim N(0, 1)$ then $X_1X_2 - X_3X_4 \sim \text{Laplace}(0, 1)$, and $(X_1^2 - X_2^2 + X_3^2 - X_4^2)/2 \sim \text{Laplace}(0, 1)$
- **Pivot:** If $X_i \sim \text{Laplace}(\mu, b)$ then $\frac{2}{b}\sum_{i=1}^{n} |X_i - \mu| \sim \chi^2_{2n}$*
- **Pivot:** If $X, Y \sim \text{Laplace}(\mu, b)$ then $\frac{|X-\mu|}{|Y-\mu|} \sim F_{2,2}$
- If $X, Y \sim \text{Unif}(0, 1)$ then $\log(\frac{X}{Y}) \sim \text{Laplace}(0, 1)$
- If $X \sim \text{EXP}(\lambda)$ and $Y \sim \text{Bernoulli}(\frac{1}{2})$ are independent, then $X(2Y - 1) \sim \text{Laplace}(0, \lambda^{-1})$
- If $X \sim \text{EXP}(\lambda)$ and $Y \sim \text{EXP}(\nu)$ are independent, then $\lambda X - \nu Y \sim \text{Laplace}(0, 1)$

2.12 F Distribution

In probability theory and statistics, the F-distribution or F-ratio, also known as Snedecor's F distribution or the Fisher–Snedecor distribution (after Ronald Fisher and George W. Snedecor), is a continuous probability distribution that arises frequently as the null distribution of a test statistic, most notably in the analysis of variance (ANOVA) and other F-tests.

1. **PDF, CDF, MGF, mean and variance** of $X \sim F_{d_1,d_2}$

   (a) **PDF**:

$$f(x) = \frac{\sqrt{\frac{(d_1x)^{d_1} d_2^{d_2}}{(d_1x+d_2)^{d_1+d_2}}}}{x\mathrm{B}\left(\frac{d_1}{2}, \frac{d_2}{2}\right)}$$

$$-\infty < x < \infty$$

   (b) **CDF**: $f(x) = I_{\frac{d_1x}{d_1x+d_2}}\left(\frac{d_1}{2}, \frac{d_2}{2}\right)$

   (c) **MGF**: DNE

   (d) **Mean and Variance**:

$$E[X] = \frac{d_2}{d_2-2} \text{ for } d_2 > 2,\ Var(X) = \frac{2d_2^2(d_1+d_2-2)}{d_1(d_2-2)^2(d_2-4)} \text{ for } d_2 > 4$$

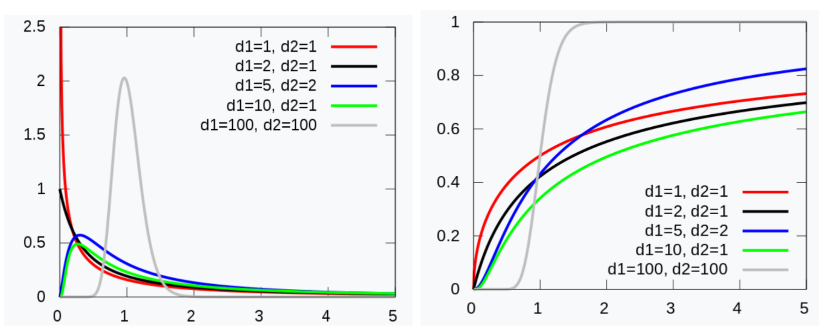

Figure 2.20: F PDF (left) CDF (right)

2. **Related Distributions**

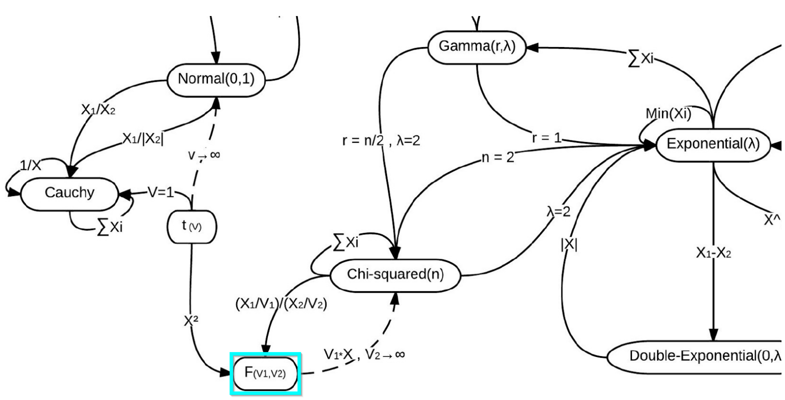

Figure 2.21: F and related distributions

- If $X \sim \chi^2_{d_1}$ and $Y \sim \chi^2_{d_2}$ are independent, then $\dfrac{X/d_1}{Y/d_2} \sim F_{d_1,d_2}$
- If $X_k \sim \Gamma(\alpha_k, \beta_k)$ are independent, then $\dfrac{\alpha_2\beta_1 X_1}{\alpha_1\beta_2 X_2} \sim F_{2\alpha_1,2\alpha_2}$
- If $X \sim \mathrm{Beta}(\frac{d_1}{2}, \frac{d_2}{2})$ then $\dfrac{d_2 X}{d_1(1-X)} \sim F$
  - Equivalently, if $X \sim F_{d_1,d_2}$ then $\dfrac{d_1 X/d_2}{1 + d_1 X/d_2} \sim \mathrm{Beta}(\frac{d_1}{2}, \frac{d_2}{2})$
- If $X \sim F_{d_1,d_2}$ then $Y = \lim_{d_2 \to \infty} d_1 X \sim \chi^2_{d1}$
- If $X \sim F_{d_1,d_2}$ then $X^{-1} \sim F_{d_2,d_1}$
- If $X \sim t_n$ then $X^2 \sim F_{1,n}$, $X^{-2} \sim F_{n,1}$
- If $X, Y \sim \mathrm{Laplace}(\mu, b)$ then $\frac{|X-\mu|}{|Y-\mu|} \sim F_{2,2}$

2.13 Student's t Distribution

In probability and statistics, Student's t-distribution (or simply the t-distribution) t_ν is a continuous probability distribution that generalizes the standard normal distribution. Like the latter, it is symmetric around zero and bell-shaped.

However, t_ν has heavier tails and the amount of probability mass in the tails is controlled by the parameter ν . For $\nu = 1$ the Student's t distribution t_ν becomes the standard Cauchy distribution, whereas for $\nu \to \infty$ it becomes the standard normal distribution $\mathrm{N}(0, 1)$.

The Student's t-distribution plays a role in a number of widely used statistical analyses, including Student's t-test for assessing the statistical significance of the difference between two sample means, the construction of confidence intervals for the difference between two population means, and in linear regression analysis.

1. **PDF, CDF, MGF, mean and variance** of $X \sim t_\nu$

   (a) **PDF**:

$$f(x) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\,\Gamma\left(\frac{\nu}{2}\right)}\left(1+\frac{x^2}{\nu}\right)^{-\frac{\nu+1}{2}}$$

$$-\infty < x < \infty$$

   (b) **MGF**: Undefined

   (c) **Mean and Variance**:

$$E[X] = 0 \text{ for } \nu > 1 \text{ otherwise undefined, } Var(X) = \tfrac{\nu}{\nu-2} \text{ for } \nu > 2,\ \infty \text{ for } 1 < \nu \leq 2$$

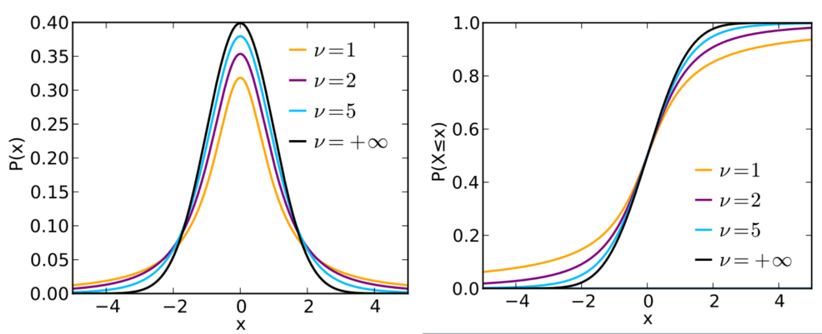

Figure 2.22: Student's t PDF (left) CDF (right)

2. **Related Distributions**

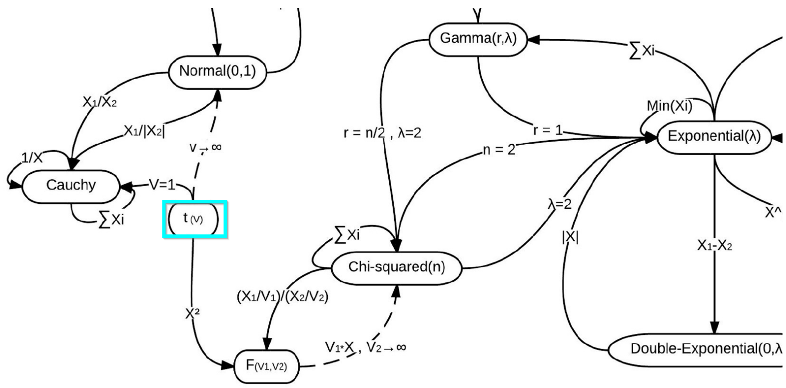

Figure 2.23: t and related distributions

- **Pivot:** If $X_1, ..., X_n \overset{i.i.d.}{\sim} N(\mu, \sigma^2)$ with sample mean $\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n}$ and sample variance $S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \bar{X})^2$ then $T = \frac{\bar{X}-\mu}{\sqrt{s^2/n}} \sim t_{n-1}$
- **Pivot**: If $Z \sim N(0,1)$ and $U \sim \chi^2_r$ are independent, then $T = \frac{Z}{\sqrt{U/r}} \sim t_r$

2.14 χ^2_n Distribution

Let us say that X is distributed χ^2_n. We know the following: A χ^2_n is the sum of the squares of n independent standard Normal r.v.s.

1. **PDF, CDF, MGF, mean and variance** of $X \sim t_\nu$

   (a) **PDF**:

   $$f(x) = \frac{1}{2^{\frac{k}{2}}\Gamma\left(\frac{k}{2}\right)} x^{\frac{k}{2}-1} e^{-\frac{x}{2}}$$
   $$0 \leq x < \infty$$

   (b) **MGF**: $(1-2t)^{-\frac{k}{2}}$ for $t < \frac{1}{2}$

   (c) **Mean and Variance**:

   $$E[X] = k,\, Var(X) = 2k$$

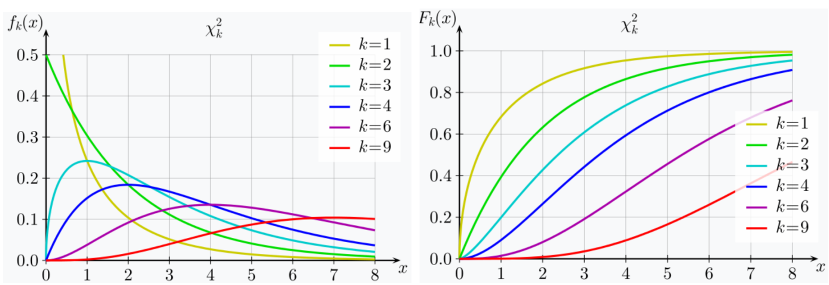

Figure 2.24: χ^2_n PDF (left) CDF (right)

2. **Related Distributions**

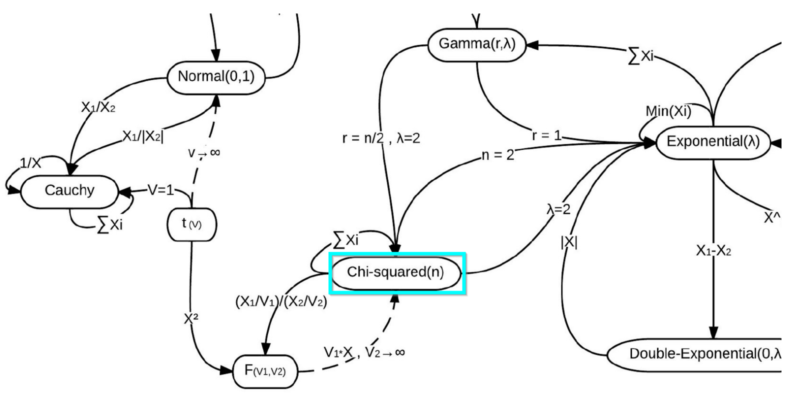

Figure 2.25: χ^2_n and related distributions

- If $Z_1, ..., Z_n \overset{i.i.d.}{\sim} N(0,1)$ then $Z_i^2 \sim \chi^2_1$ and $Q = \sum_{i=1}^{n} Z_i^2 \sim \chi^2_n$
- If $X \sim \chi^2_k$ then as $k \to \infty$, $\frac{(X-k)}{\sqrt{2k}} \xrightarrow{d} N(0,1)$ (CLT)
- If $X \sim \text{Gamma}(\alpha, \beta)$, for $\alpha = \frac{n}{2}$, $\beta = 2$, $X \sim \chi^2_n$

- If $X \sim \text{EXP }(\theta), \sum\limits_{i=1}^{n} X_i \sim \text{Gamma}(n, \theta)$ and $T = \frac{2 \sum\limits_{i=1}^{n} X_i}{\theta} \sim \chi^2_{2n}$
- If $X \sim F_{d_1,d_2}$ then $Y = \lim\limits_{d_2 \to \infty} d_1 X \sim \chi^2_{d1}$
- If $X \sim \chi^2_k$ and $c > 0$ then $cX \sim \text{Gamma}(\frac{k}{2}, 2c)$
- If $X \sim \chi^2_2$ then $X \sim \text{EXP}(\frac{1}{2})$
- If $X \sim \chi^2_{\nu_1}$ and $Y \sim \chi^2_{\nu_2}$ are independent, then $\frac{X}{X+Y} \sim \text{Beta}(\frac{\nu_1}{2}, \frac{\nu_2}{2})$
- If $X \sim \text{Unif}(0.1)$ then $-2\log(X) \sim \chi^2_2$
- If $X_i \sim \text{Laplace}(\mu, b)$ then $\frac{2}{b} \sum_{i=1}^{n} |X_i - \mu| \sim \chi^2_{2n}$

2.15 Irwin-Hall

A random variable with Irwin-Hall distribution is defined as the sum of a number of independent random variables, each having a uniform distribution. For this reason it is also known as the uniform sum distribution. So, if $U_i \sim \text{Unif}(0,1)$ and $U_1, ..., U_n$ are *i.i.d.* then $X = \sum_{i=1}^{n} U_i \sim \text{IrwinHall}(n)$

1. **PDF, CDF, MGF, mean and variance** of $X \sim \text{IrwinHall}(n)$

   (a) **PDF**:

   $$f(x) = \frac{1}{(n-1)!} \sum_{k=1}^{\lfloor x \rfloor} (-1)^k \binom{n}{k} (x-k)_+^{n-1}$$

   $$\text{where } (x-k)_+ = \begin{cases} x-k & x-k \geq 0 \\ 0 & x-k < 0 \end{cases}$$

   $$0 \leq x \leq n$$

   $$n \in \mathbb{N}$$

   (b) **CDF**: $F(x) = \frac{1}{n!} \sum_{k=1}^{\lfloor x \rfloor} (-1)^k \binom{n}{k} (x-k)_+^n$

   (c) **MGF**: $M_X(t) = \left(\frac{e^t - 1}{t}\right)^n$

   (d) **Mean and Variance**:

   $$E[X] = \frac{n}{2},\ Var(X) = \frac{n}{12}$$

2. **Example problems** and key steps