Bhoris Dhanjal

Probability and Sampling Distributions (B)

Lecture Notes
for SSTA401

Contents

Chapter 1

Transformation of random variables & standard univariate continuous probability distributions

1.1 Uniform/Rectangular distributions

Definition 1.1. *A r.v. X is said to follow uniform distribution over an interval (a, b) if its pdf is constant over the entire range.*

1.1.1 PDF of uniform distribution

Theorem 1.2. *PDF of uniform distribution*

$$P(x) = k \qquad\qquad a < x < b$$
$$= 0 \qquad\qquad otherwise$$

- $\int_a^b f(x)\, dx = \int_a^b k\, dx = k[x]_a^b = k(b-a) = 1$, therefore $k = \frac{1}{b-a}$
- We denote it as, $X \sim U(a, b)$
- $f(x) = \frac{1}{b-a}$
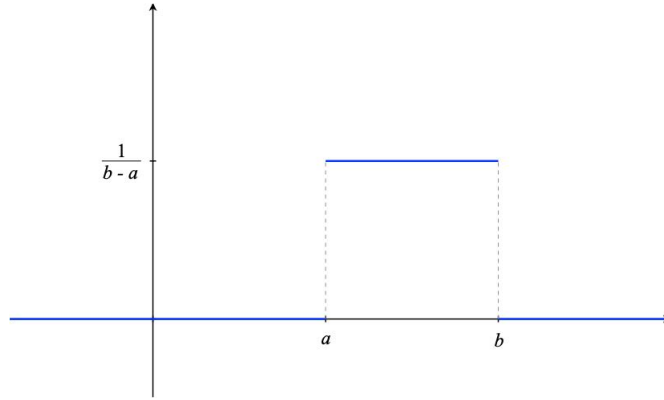
1.1.2 CDF of uniform distribution

Theorem 1.3. *CDF of uniform distribution*

$$
\begin{aligned}
F(x) &= 0 & x \leq a \\
&= P(X \leq x) = \int_a^x f(x)\,dx = \frac{x-a}{b-a} & a < x < b \\
&= 1 & x \geq b
\end{aligned}
$$

1.1.3 Expectation and variance of uniform distribution

Theorem 1.4. *Expected value of $X \sim U(a,b)$ is equal to $\frac{(a+b)}{2}$*

Proof. Consider the expectation of the uniform distribution as,

$$
\begin{aligned}
E[x] &= \int_a^b x P(x)\,dx \\
&= \int_a^b x \frac{1}{b-a}\,dx \\
&= \frac{1}{b-a} \int_a^b x\,dx \\
&= \frac{a+b}{2}
\end{aligned}
$$

$\square$

Theorem 1.5. *Variance of uniform distribution is equal to $\frac{1}{12}(b-a)^2$*

Proof. We begin by finding out $E[X^2]$

$$E[X^2] = \int_a^b x^2 P(x)\,dx$$

$$= \int_a^b x^2 \frac{1}{b-a}\,dx$$

$$= \frac{1}{b-a} \int_a^b x^2\,dx$$

$$= \frac{1}{3}\left(a^2 + ab + b^2\right)$$

Now we can find the variance as $V[X] = E[X^2] - E[X]^2$ as follows,

$$V[X] = E[X^2] - E[X]^2$$

$$= \frac{1}{3}\left(a^2 + ab + b^2\right) - \left(\frac{a+b}{2}\right)^2$$

$$= \frac{(b-a)^2}{12}$$

$\square$

1.1.4 Raw moments of uniform distribution

The rth raw moment of the uniform distribution is given as

$$\mu_r' = E[X^r] = \int_a^b x^r f(x)\,dx$$

$$= \frac{1}{b-a}\left[\frac{x^{r+1}}{r+1}\right]_a^b = \frac{b^{r+1} - a^{r+1}}{(b-a)(r+1)}$$

Example 1.6. *Suppose in a quiz there are 30 participants. A question is given to all 30 participants and the time allowed is 25 seconds.*

Proof. Let X denote the time to respond.
$X \sim U(0, 25)$, the pdf is given by $f(x) = \frac{1}{25}; 0 < x < 25$ and 0 otherwise.

$$P(x \leq 6) = \int_0^6 f(x)\,dx = \int_0^6 \frac{1}{25}\,dx = \frac{151}{25}$$

$$P(6 \leq x \leq 10) = \int_6^1 0 f(x)\,dx = \int_6^{10} \frac{1}{25}\,dx = \frac{101}{25}$$

$\square$

Example 1.7. *A r.v. x is said to follow uniform dist with $\mu = 1$ and $V(x) = 4/3$. Obtain $P(x < 0)$.*

Proof. First begin by finding out the parameters for the unfirom distribution. First consider the mean,

$$\mu = 1$$
$$\frac{a+b}{2} = 1$$
$$a + b = 2$$

Then consider the variance,

$$V(x) = \frac{4}{3}$$
$$\frac{(b-a)^2}{12} = \frac{4}{3}$$
$$(b-a)^2 = 16$$

Solving two simultaneous equations we get $a = -1, b = 3$.
Therefore, we have $X \sim U(-1, 3)$

$$P(x \leq 0) = F(0) = \frac{0+1}{4} = \frac{1}{4}$$

$\square$

Example 1.8. *If $X \sim U(-3, 3)$, find $P(x < 2)$, $P(|x| < 2)$, $P(|x - 2| < 2)$, also obtain k if $P(x > k) = 1/3$*

Proof.

$$P(x < 2) = F(2) = \frac{2+3}{6} = \frac{5}{6}$$
$$P(|x| < 2) = \int_{-} 2^2\frac{1}{6}\, dx = \frac{2}{3}$$
$$P(|x - 2| < 2) = \int_0^3 \frac{1}{6} = \frac{1}{2}$$
$$P(x > k) = 1/3 \implies \ldots$$

Complete this l8r alig8r

$\square$

1.1.5 MGF of Uniform distribution

Theorem 1.9. *MGF of Uniform distribution* $= \frac{e^{bt}-e^{at}}{t(b-a)}, t \neq 0$ *and* $= 1, t = 0$

Proof.

$$M_x(t) = E[e^{tx}] = \int_a^b \frac{e^{tx}}{b-a}\,dt = \frac{e^{bt}-e^{at}}{(b-a)t}$$

The Taylor series for this can be expressed as the following,

$$M_x(t) = \frac{b-a}{b-a} + \frac{b^2-a^2}{2(b-a)}t + \frac{b^3-a^3}{3(b-a)}\frac{t^2}{2!} + \cdots$$

Therefore we can say,

$$\mu_1' = \text{coeff of } t = \frac{b^2-a^2}{2(b-a)} = \frac{a+b}{2}$$

$$\mu_2' = \text{coeff of } \frac{t^2}{2!} = \frac{b^3-a^3}{3(b-a)}$$

And we can say $\mu_2 = \ldots$ $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

1.1.6 Applications of uniform distribution

1. Assumption of uniform death for insurance
   $\vdots$

Write sumthin here

1.2 Gamma distribution

Definition 1.10 (Gamma distribution). *A r.v. 'X' is said to follow gamma distribution* $X \sim G(\lambda, \theta)$. *Where* $\lambda = $ *shape parameter and* $\theta = $ *scale parameter.*

1.2.1 PDF of Gamma distribution

Definition 1.11 (PDF of Gamma distribution).

$$f(x, \lambda, \theta) = \frac{\theta^\lambda}{\Gamma(\lambda)} e^{-\theta x} x^{\lambda-1} \qquad\qquad x > 0, \lambda > 0, \theta > 0$$

$$= 0 \qquad\qquad\qquad otherwise$$

Where $\Gamma(\lambda) = (\lambda-1)! = (\lambda-1)\Gamma(\lambda-1)$.

Corollary 1.12. *If $\theta = 1$ we will have gamma distribution with a single parameter λ which is called the standard gamma distribution.*

$$X \sim G(\lambda) = \frac{e^{-x}x^{\lambda-1}}{\Gamma(\lambda)} \qquad\qquad x > 0, \lambda > 0$$
$$= 0 \qquad\qquad otherwise$$

Corollary 1.13. *If $\lambda = 1, X \sim G(1,\theta) = Exp(\theta)$.*

Corollary 1.14. *If $\lambda = 1, \theta = 1$, $X \sim$ Standard exponential distribution, i.e.*

$$f(x) = e^{-x} \qquad\qquad x > 0$$
$$= 0 \qquad\qquad otherwise$$

Definition 1.15 (Gamma function)**.**

$$\Gamma(\lambda) = \int_0^\infty e^{-x}x^{\lambda-1}\,dx$$

Definition 1.16 (Gamma integral)**.**

$$\int_0^\infty e^{-\theta x}x^{\lambda-1}\,dx = \frac{\Gamma(\lambda)}{\theta^\lambda}$$

1.2.2   CDF of Gamma distribution

Theorem 1.17. *CDF of Gamma distribution is given as*

$$F(x) =$$

Proof.

$$F(x) = P(X < x) = \int_0^x \frac{\theta^\lambda e^{-\theta x}x^{\lambda-1}}{\Gamma(\lambda)}\,dx$$
$$= \frac{\theta^\lambda}{\Gamma(\lambda)}\int_0^x x^{\lambda-1}e^{-\theta x}\,dx$$

$\square$

1.2.3 Raw moments of Gamma distribution

Theorem 1.18. *The rth raw moment of the Gamma distribution is given by*

$$\mu'_r = \frac{\Gamma(\lambda + r)}{\Gamma(\lambda)\theta^r}$$

Proof.

$$\mu'_r = E[x^r] = \int_0^\infty \frac{x^r \theta^\lambda e^{-\theta x} x^{\lambda-1}}{\Gamma(\lambda)} \, dx$$
$$= \int_0^\infty \frac{\theta^\lambda e^{-\theta x} x^{\lambda+r-1}}{\Gamma(\lambda)} \, dx$$
$$= \frac{\Gamma(\lambda + r)}{\Gamma(\lambda)\theta^r}$$

$\square$

1.2.4 Mean and Variance of Gamma distribution

Now we can find μ'_1, μ'_2

$$E[x] = \mu'_1 = \frac{\lambda}{\theta}$$
$$\mu'_2 = \frac{\lambda(\lambda + 1)}{\theta^2}$$
$$V[x] = \mu_2 = \mu'_2 - \mu'^2_1 = \frac{\lambda(\lambda + 1)}{\theta^2} - \frac{\lambda^2}{\theta^2} = \frac{\lambda}{\theta^2}$$

1.2.5 MGF of Gamma distribution

$$E[e^{tx}] = \int_0^\infty e^{tx} \frac{\theta^\lambda e^{-\theta x} x^{\lambda-1}}{\Gamma(\lambda)} \, dx$$
$$= \frac{\theta^\lambda}{\Gamma(\lambda)} \int_0^\infty e^{-(\theta-t)x} x^{\lambda-1} \, dx$$
$$= \frac{\theta^\lambda}{\Gamma(\lambda)} \frac{\Gamma(\lambda)}{(\theta - t)^\lambda} = \left(\frac{\theta}{\theta - t}\right)^\lambda$$
$$= \left(1 - \frac{t}{\theta}\right)^{-\lambda}$$

1.2.6 CGF of Gamma distribution

$$K_x(t) = \log\left(1 - \frac{t}{\theta}\right)^{-\lambda}$$

$$= -\lambda \log\left(1 - \frac{t}{\theta}\right)$$

$$= \frac{\lambda t}{\theta} + \frac{\lambda t^2}{2\theta^2} + \frac{\lambda t^3}{3\theta^3} + \cdots$$

Using this we can get the mean and variance easily.

$$\text{Mean} = k_1 = \frac{\lambda}{\theta}$$

$$\text{Variance} = k_2 = \frac{\lambda}{\theta^2}$$

1.2.7 Additive property of Gamma distribution

If $X_i(i = 1, \ldots, k)$ are k independent Gamma distributions with parameters $\lambda_1, \lambda_2, \ldots, \lambda_k$ and θ respectively, then,

$$\sum_{i=1}^{k} X_i \sim G\left(\sum_{i=1}^{k} \lambda_i, \theta\right)$$

$$M_{X_i}(t) = \left(1 - \frac{t}{\theta}\right)^{-\lambda_i}$$

Let $Z = \sum X_i$

$$M_Z(t) = \prod_{i=1}^{k} \left(1 - \frac{t}{\theta}\right)^{-\lambda_i}$$

$$= \left(1 - \frac{t}{\theta}\right)^{-\sum \lambda_i}$$

By uniqueness property of mgf

$$\sum_i X_i \sim G\left(\sum_i \lambda_i, \theta\right)$$

1.2.8 Limiting form of Gamma distribution

$$\beta_1 = \frac{4}{\lambda}, \text{ as } \lambda \to \infty, \beta_1 \to 0 \implies \text{Normal dist}$$

$$\beta_2 = 3 + \frac{6}{\lambda} \text{ as } \lambda \to \infty, \beta_2 \to 3 \implies \text{Normal dist}$$

Note that they are both independent of θ.
Therefore, as $\lambda \to \infty$ we have $G(\lambda, \infty) \to N\left(\frac{\lambda}{\theta}, \frac{\lambda}{\theta^2}\right)$.

1.2.9 Applications of Gamma distribution

Idk write something bruh

1.3 Exponential distribution

1.3.1 PDF of Exponential Distribution

Definition 1.19 (PDF of Exponential distribution)**.** *A r.v. x is said to follow the exponential distribution with parameter θ if its pdf is given by*

$$f(x) = \theta e^{-\theta x} \qquad\qquad x \geq 0, \theta > 0$$
$$= 0 \qquad\qquad otherwise$$

1.3.2 INCOMPLETE CDF of exponential distribution

$$F[x] = 1 - e^{-\theta x}$$

FILL THIS UP

1.3.3 Raw moment of exponential distribution

Theorem 1.20. *The rth raw moment for exponential distribution is given by*

$$\mu_r' = \frac{r!}{\theta^r}$$

Proof.

$$\mu_r' = E[x^r] = \int_0^\infty x^r \theta e^{-\theta x} \, dx$$
$$= \frac{\Gamma(r+1)}{\theta^r}$$
$$= \frac{r!}{\theta^r}$$

$\square$

1.3.4 Mean and variance of exponential distribution

Theorem 1.21. *The mean of exponential distribution is given be*

$$\mu = \frac{1}{\theta}$$

Proof. Consider $r = 1$,

$$\mu_1' = \frac{1}{\theta}$$

$\square$

Theorem 1.22. *The variance of the exponential distribution is given by*

$$\mu_2 = \frac{1}{\theta^2}$$

Proof. First find μ_2'

$$\mu_2' = \frac{2}{\theta^2}$$

So now we can compute the variance as $\frac{1}{\theta^2}$

$\square$

1.3.5 MGF of exponential distribution

Theorem 1.23. *MGF of exponential distribution is given by*

$$M_x(t) = \left(1 - \frac{t}{\theta}\right)^{-1}$$

Proof.

$$M_x(t) = E[e^{tx}]$$
$$= \int_0^\infty e^{tx}\theta e^{-\theta x}\, dx$$
$$= \theta \int_0^\infty e^{x(t-\theta)}x^{1-1}\, dx$$
$$= \frac{\theta\Gamma(1)}{\theta - t}$$
$$= \frac{\theta}{\theta - t}$$
$$= \left(1 - \frac{t}{\theta}\right)^{-1}$$

$\square$

1.3.6 CGF of exponential distribution

Theorem 1.24. *CGF of exponential distribution is given by*

$$K_x(t) = -\log\left(1 - \frac{t}{\theta}\right)$$

Proof.

$$K_x(t) = \log\left(1 - \frac{t}{\theta}\right)^{-1}$$

$$= -\log\left(1 - \frac{t}{\theta}\right)$$

$$= \frac{t}{\theta} + \frac{t^2}{2\theta^2} + \frac{t^3}{3\theta^3}$$

We can say the general r^{th} cumulant is given by $K_r = \frac{(r-1)!}{\theta^r}$ $\square$

1.3.7 Additive property of exponential variates

Theorem 1.25. *If $x_1, x_2, \ldots, x_k$ are k independent exponential variates each with parameter θ then*

$$\sum_{i=1}^{k} x_i \sim G(k, \theta)$$

Proof. We will do this with the MGF. Consider taht $Z = \sum i = 1^k x_i$.

$$M_z(t) = \prod_{i=1}^{k} M_x(t)$$

$$= \prod_{i=1}^{k} \left(1 - \frac{t}{\theta}\right)^{-1}$$

$$= \left(1 - \frac{t}{\theta}\right)^{-k}$$

Therefore, (by uniqueness property of MGF) comparing this MGF to that of the gamma distribution we can say that,

$$\sum_{i=1}^{k} x_i = Z \sim G(k, \theta)$$

$\square$

1.3.8   Lack of memory of exponential distribution

Theorem 1.26. *For a exponentially distributed random variate, $P[x > a+b \mid x > a] = P[x > b]$*

Proof. Let $X \sim E(\theta)$. Consider first case

$$
\begin{aligned}
P[x > a + b \mid x > a] &= \frac{P[x > a + b]}{P[x > a]} \\
&= \frac{\int_{a+b}^{\infty} \theta e^{-\theta x} \, dx}{\int_{a}^{\infty} \theta e^{-\theta x} \, dx} \\
&= \frac{e^{-\theta a + b}}{e^{-\theta a}} \\
&= e^{-\theta b}
\end{aligned}
$$

Consider second case now,

$$
P[x > b] = \int_{b}^{\infty} \theta e^{-\theta x} \, dx = e^{-\theta b}
$$

Equality holds. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad \square$

1.4   INCOMPLETE Laplace distribution (Double exponential)

1.4.1   PDF

Definition 1.27 (PDF of Laplace distribution). $X \sim L(\lambda, \mu)$

$$
f(x) = \begin{cases} \frac{1}{2\lambda} e^{-\left|\frac{x-\mu}{\lambda}\right|} & -\infty < x < \infty \\ 0 & otherwise \end{cases}
$$

1.4.2   CDF

Definition 1.28 (CDF of Laplace distribution).

$$
F[x] = \Big\{ content...
$$

1.4.3   Raw moment

Theorem 1.29. *The rth raw moment for the Laplace distribution is given by*

$$\mu_r' =$$

Proof.

$$\mu_r' = E[x^r] = \int_{-\infty}^{\infty} \frac{x^r}{2\lambda} e^{-\left|\frac{x-\mu}{\lambda}\right| dx|}$$

Transform $(x - \mu)/\lambda = z$

$$= \frac{1}{2\lambda} \left(\int_{-\infty}^{\infty} (z\lambda + \mu)^r e^{-|z|} \lambda \, dz \right)$$

$$= \frac{1}{2} \left(\int_{-\infty}^{\infty} \sum_{k=0}^{r} \binom{r}{k} (z - \lambda)^k \mu^{r-k} e^{-|z| \, dz} \right)$$

$$= \frac{1}{2} \sum_{k=0}^{r} \left[\binom{r}{k} \lambda^k \mu^{r-k} \int_{-\infty}^{\infty} z^k e^{-|z|} \, dz \right]$$

Complete this up

$$= \frac{1}{2} \sum_{k=0}^{r} \left[\binom{r}{k} \lambda^k \mu^{r-k} k! (1 + (-1)^k) \right]$$

$\square$

1.4.4   Mean and variance

We can do this with the raw moments above but instead we will do it with the PDF.

Theorem 1.30. *Expectation of laplace distribution is given as*

$$E[x] =$$

Proof.

$$E[x] = \int_{-\infty}^{\infty} x f(x) \, dx$$

$$= \int_{-\infty}^{\infty} \frac{x}{2\lambda} e^{-\left|\frac{x-\mu}{\lambda}\right|} \, dx$$

Split it around μ

$$= \frac{1}{2\lambda} \left(\int_{-\infty}^{\mu} x e^{\frac{x-\mu}{\lambda}} \, dx + \int_{\mu}^{\infty} x e^{-\frac{x-\mu}{\lambda}} \, dx \right)$$

$$= \frac{1}{2\lambda} \left[e^{-\mu/\lambda} \int_{-\infty}^{\mu} x e^{x/\lambda} \, dx + e^{\mu/\lambda} \int_{\mu}^{\infty} x e^{-x/\lambda} \, dz \right]$$

$$= \frac{1}{2\lambda} \left[e^{-\mu/\lambda} \lambda (x-\lambda) e^{x/\lambda} - e^{\mu/\lambda} (\lambda(x+\lambda) e^{-x/\lambda}) \right]$$

$$= \mu$$

$\square$

Theorem 1.31. *Expectation of x^2 in Laplace distribution is given be*

$$E[x^2] = bruh$$

Proof.

$$E[x^2] = \int_{-\infty}^{\infty} x^2 \frac{1}{2\lambda} e^{-\left| \frac{x-\mu}{\lambda} \right|}$$

Split it around μ

$$= \frac{1}{2\lambda} \left(\int_{-\infty}^{\mu} x^2 e^{\frac{x-\mu}{\lambda}} \, dx + \int_{\mu}^{\infty} x^2 e^{-\frac{x-\mu}{\lambda}} \, dx \right)$$

$$= \frac{1}{2\lambda} \left[e^{-\mu/\lambda} \int_{-\infty}^{\mu} x^2 e^{x/\lambda} \, dx + e^{\mu/\lambda} \int_{\mu}^{\infty} x^2 e^{-x/\lambda} \, dx \right]$$

$$= \frac{1}{2\lambda} \left[e^{-\mu/\lambda} (\lambda(x^2 - 2\lambda x + 2\lambda^2) e^{x/\lambda}) - e^{\mu/\lambda} (\lambda(x^2 + 2\lambda x + 2\lambda^2) e^{-x/\lambda}) \right]$$

$$= 2\lambda^2$$

$\square$

Theorem 1.32. *Variance of Laplace distribution is given as*

$$V[x] =$$

1.4.5 MGF

Theorem 1.33. *MGF of the Laplace distribution is given by*

$$M_x(t) = bruh$$

Proof.

$$M_x(t) = E[e^{tx}] = \int_{-\infty}^{\infty} \frac{1}{2\lambda} e^{tx - \left|\frac{x-\mu}{\lambda}\right|}$$

$$= \frac{1}{2\lambda} \left[e^{-\mu/\lambda} \int_{-\infty}^{\mu} e^{x\left(t + \frac{1}{\lambda}\right)} \, dx + e^{\mu/\lambda} \int_{\mu}^{\infty} e^{-x\left(\frac{1}{\lambda} - t\right)} \, dx \right]$$

$$= \frac{1}{2\lambda} \left[e^{-\mu/\lambda} \left(\frac{e^{\mu\left(\frac{1}{\lambda} + t\right)}}{\frac{1}{\lambda} + t} \right) + e^{\mu/\lambda} \left(\frac{-e^{\mu\left(\frac{1}{\lambda} - t\right)}}{-\frac{1}{\lambda} + t} \right) \right]$$

$$= \frac{1}{2\lambda} \left[\frac{e^{\mu t}}{t + \frac{1}{\lambda}} - \frac{e^{\mu t}}{t - \frac{1}{\lambda}} \right]$$

$\square$

Plot a graph for the beta-1 dsitribution when alpha=5, beta=2

1.4.6   CGF

1.5   Beta distribution of Type-I

1.5.1   PDF

Definition 1.34 (PDF of Beta I).

$$f(x) = \begin{cases} \frac{1}{\beta(m,n)} x^{m-1}(1-x)^{n-1} & 0 < x < 1; m, n > 0 \\ 0 & otherwise \end{cases}$$

Where $\beta(m,n) = \frac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)}$

Note the following,

1. We can say, $X \sim \beta_1(m, n)$ where m, n are the parameters of the distribution.

2. Since $f(x)$ is a pdf we have the following,

$$\int_0^1 f(x) \, dx = \int_0^1 \frac{1}{\beta(m, n)} x^{m-1}(1-x)^{n-1} \, dx$$

$$= \int_0^1$$

1.5.2 Raw moments

Theorem 1.35. *The rth raw moment of the Beta I distribution is given by*

$$\mu'_r = \frac{\Gamma(m+n)\Gamma(r+m)}{\Gamma(m)\Gamma(m+n+r)}$$

Proof.

$$\mu'_r = E[x^r] = \int_0^1 \frac{1}{\beta(m,n)} x^{r+m-1}(1-x)^{n-1}\, dx$$
$$= \frac{\Gamma(m+n)\Gamma(r+m)}{\Gamma(m)\Gamma(m+n+r)}$$

$\square$

1.5.3 Mean and Variance

Theorem 1.36. *Mean of Beta I distribution is given by*

$$E[x] = \frac{m}{m+n}$$

Proof.

$$E[x] = \mu'_1 = \frac{\Gamma(m+n)\Gamma(m+1)}{\Gamma(m)+\Gamma(m+n+1)} = \frac{m}{m+n}$$

$\square$

Theorem 1.37. *Variance of Beta I distribution is given by*

$$V[x] = \frac{mn}{(m+n)^2(m+n+1)}$$

Proof.

$$\mu'_2 = \frac{(m+1)(m)}{(m+n)(m+n+1)}$$

So now we have the variance given as,

$$\mu_2 = \mu'_2 - \mu'^2_1$$
$$= \frac{mn}{(m+n)^2(m+n+1)}$$

$\square$

1.6 Beta distribution of Type-II

1.6.1 PDF

Definition 1.38 (PDF of Beta-II distribution).

$$f(x) = \begin{cases} \frac{1}{\beta(m,n)} \frac{x^{m-1}}{(1+x)^{m+n}} & 0 < x < \infty; m, n > 0 \\ 0 & otherwise \end{cases}$$

Note the following,

1. X is said to follow $\beta_2(m,n)$ as $X \sim \beta_2(m,n)$

2.

$$\int_0^\infty f(x)\,dx = \int_0^\infty \frac{x^{m+1}}{(1+x)^{m+n}} = \beta(m,n)$$

1.6.2 Raw moments

Theorem 1.39 (Raw moments of Beta-2 distribution). *The raw moments of the Beta-2 distribution is given by*

$$\mu_r' = \frac{\Gamma(m+r)\Gamma(n-r)}{\Gamma(m)\Gamma(n)}$$

Proof.

$$\mu_r' = E[x^r] = \int_0^\infty \frac{1}{\beta(m,n)} \frac{x^{m+r-1}}{(1+x)^{m+n}}\,dx$$
$$= \frac{\Gamma(m+r)\Gamma(n-r)}{\Gamma(m)\Gamma(n)}$$

$\square$

1.6.3 Mean and variance

Theorem 1.40 (Mean of Beta-2 distribution). *The mean of Beta-2 distribution is given by*

$$E[x] = \frac{m}{n-1}$$

Proof.

$$E[x] = \mu'_1 = \frac{\Gamma(m+1)\Gamma(n-1)}{\Gamma(m)\Gamma(n)}$$

$$= \frac{m}{n-1}$$

□

Theorem 1.41 (Variance of Beta-2 distribution)**.** *The variance of Beta-2 distribution is given by*

$$V[x] = \frac{m(m+n-1)}{(n-2)(n-1)^2}$$

Proof. First consider the 2nd raw moment,

$$\mu'_2 = \frac{m(m+1)}{(n-2)(n-2)}$$

Now we can compute the variance as follows

$$V[x] = \mu_2 = \mu'_2 - \mu'^2_1 = \frac{m(m+n-1)}{(n-2)(n-1)^2}$$

□

1.7 Transformation of variables

1.7.1 One dimensional random variable

Let X be a continuous random variable with pdf $f(x)$ and let $Y = g(x)$ be a strictly monotonic function of X with unique inverse.
Assume that $g(x)$ is differentiable and is continuous for all x, then the pdf of r.v. Y is given by

$$h(y) = f(x) \cdot \det \left| \frac{dx}{dy} \right| = \left| \frac{dx}{dy} \right|$$

where r.v. x is expressed in terms of y. Steps to solve,

1. Write pdf of r.v. X.

2. Express old variable X in terms of new variable Y.

3. Write the range of the new variable.

4. Obtain J where $J = \left| \frac{dx}{dy} \right|$ and $|J|$.

5. Obtain $h(y) = f(x) \cdot |J|$, where X is expressed in terms of Y.

Remark 1.42. *For $2-1$ correspondence, i.e. for ever 2 values of X is there is only one value of Y, then multiple $|J|$ with 2.*

Remark 1.43.

For $1-2$ correspondence i.e., for every 1 value of x if there are 2 values of Y then multiply $|J|$ with $\frac{1}{2}$.

Example 1.44. *If a r.v. $X \sim B_1(m, n)$ obtain the distribution of $Y = 1 - X$.*

Proof. First begin by stating the pdf of X.

$$f(x) = \begin{cases} \frac{1}{\beta(m,n)} x^{m-1} (1-x)^{n-1} & 0 < x < 1; m, n > 0 \\ 0 & \text{otherwise} \end{cases}$$

Now $X = 1 - Y$ this ranges from $1 - Y = 0$ to $1 - Y = 1$. So $0 < Y < 1$ again.
Now compute J

$$\frac{dx}{dy} = \frac{1}{dy}(1 - y)$$
$$J = -1$$
$$|J| = 1$$

We multiply this with $f(x)$ to get $h(y)$.

$$h(y) = f(x) \cdot |J|$$
$$h(y) = f(x)$$

So $h(y) \sim B(n, m)$. The order changes. $\qquad \square$

Example 1.45. *A r.v. $X \sim B_2(m, n)$. Obtain the distribution of Y where $Y = \frac{1}{1+X}$.*

Proof. First state the pdf,

$$f(x) = \begin{cases} \frac{1}{\beta(m,n)} \frac{x^{m-1}}{(1+x)^{m+n}} & 0 < x < \infty; m, n > 0 \\ 0 & \text{otherwise} \end{cases}$$

Now state X in terms of Y, we have $X = \frac{1}{Y} - 1$.

Compute the new ranges now we have $\frac{1}{Y} - 1 = 0$ so $Y = 1$ as one side then $\frac{1}{Y} - 1 = \lim_{m \to \infty} m$ so to $Y = 0$.

The new ranges are $0 < Y < 1$. Now compute $|J|$,

$$J = \frac{dx}{dy} = \frac{1}{dy}\left(\frac{1}{y} - 1\right)$$

$$= -\frac{1}{y^2}$$

$$|J| = \frac{1}{y^2}$$

So now we can compute $h(y)$ as follows,

$$h(y) = f(x)|J|$$

$$= \frac{1}{\beta(m,n)} \frac{\left(\frac{1}{y} - 1\right)^{m-1}}{(1/y)^{m+n}} \frac{1}{y^2}$$

$$= \frac{1}{\beta(m,n)} y^{n-1}(1-y)^{m-1}$$

This is for the range we have and 0 otherwise. But I'm too lazy to typeset that out as a cases.

So we now have $Y \sim B_1(n, m)$. □

1.8 Two dimensional r.v.

Let X and Y be two continuous independent r.v. with joint pdf $f(x, y)$. Say $U = g(x, y)$ and $V = h(x, y)$ are two other r.v. then the joint pdf of U and V is given by,

$$k(u, v) = f(x, y) \cdot |J|$$

where X, Y are expressed in terms of U, V. Here we have the Jacobian as follows,

$$\begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial u}{\partial u} & \frac{\partial y}{\partial v} \end{bmatrix} = \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial y}{\partial u} \\ \frac{\partial x}{\partial v} & \frac{\partial y}{\partial u} \end{bmatrix}$$

1.8.1 Steps to solve

1. Write the pdf of X and Y, i.e. $f(x, y)$.

2. Express old variable in terms of new variable.

3. Obtain range of the new variable.

4. Obtain J and $|J|$.

5. Obtain $k(u, v) = f(x, y)|J|$.

Example 1.46. *X and Y are two independent gamma variates with parameters a and b respectively.*

1. *Obtain the joint distribution of u and v where $u = x + y, v = \frac{x}{x+y}$.*

2. *Show that u, v are independent and identify their distributions.*

Proof. Consider the pdf of gamma function as follows,

$$X \sim G(\lambda) = \frac{e^{-x}x^{\lambda-1}}{\Gamma(\lambda)} \qquad\qquad x > 0, \lambda > 0$$
$$= 0 \qquad\qquad \text{otherwise}$$

Where $\Gamma(\lambda) = (\lambda - 1)! = (\lambda - 1)\Gamma(\lambda - 1)$.

$$f_1(x) = \frac{1}{\Gamma(a)}e^{-x}x^{a-1}$$
$$f_2(x) = \frac{1}{\Gamma(b)}e^{-x}x^{b-1}$$

Find $f(x, y) = f_1(x)f_2(y)$

$$f(x, y) = \frac{1}{\Gamma(a)\Gamma(b)}e^{-x-y}x^{a-1}x^{b-1} \qquad\qquad x, y, a, b, > 0$$
$$= 0 \qquad\qquad \text{otherwise}$$

We now have the new variables U, V $U = X + Y, V = \frac{X}{X+Y}$. This implies that $X = UV, Y = U(1 - V)$.
We need to find the new ranges now. Since $X, Y > 0$ we have $U > 0$ and $X < X + Y \implies \frac{x}{x+y} < 1 \implies v < 1$. And $0 < V < 1$.

Find the Jacobian,

$$J = \begin{bmatrix} v & u \\ 1 - v & -u \end{bmatrix} = -u$$
$$|J| = u$$

The joint distribution is then given as,

$$k(u,v) = \frac{1}{\Gamma(a)\Gamma(b)} e^{-(uv+u-uv)} (uv)^{a-1}[u(1-v)]^{b-1} \cdot u$$

$$= \frac{1}{\Gamma(a)\Gamma(b)} e^{-u} u^{a-1+b-1+1} v^{a-1}(1-v)^{b-1} \times \frac{\Gamma(a+b)}{\Gamma(a+b)}$$

$$= \frac{1}{\Gamma(a+b)} e^{-u} u^{a+b-1} \times \frac{1}{\beta(a,b)} v^{a-1}(1-v)^{b-1}$$

$$= k_1(u)k_2(v)$$

So u and v are independent r.v. and $U \sim G(a+b), V \sim \beta_1(a,b)$ $\qquad\square$

Example 1.47. *X and Y are two independent r.v. $X \sim G(a)$ and $Y \sim G(b)$. We have $U = X + Y$ and $W = \frac{X}{Y}$. Show that U, W are independent and identify the distribution.*

Proof. We know the following,

$$f_1(x) = \frac{e^{-x}x^{a-1}}{\Gamma(a)} \qquad\qquad x > 0, a > 0$$

$$= 0 \qquad\qquad\qquad \text{otherwise}$$

and,

$$f_2(y) = \frac{e^{-y}y^{b-1}}{\Gamma(b)} \qquad\qquad x > 0, b > 0$$

$$= 0 \qquad\qquad\qquad \text{otherwise}$$

Now the joint distribution $f(x,y)$ is given by its product since they are independent,

$$f(x,y) = \frac{e^{-x}x^{a-1}}{\Gamma(a)} \times \frac{e^{-y}y^{b-1}}{\Gamma(b)} \qquad x > 0, y > 0; a, b > 0$$

$$= 0 \qquad\qquad\qquad\qquad\qquad \text{otherwise}$$

Now we compute the new ranges $X = \frac{UW}{W+1}$ and $Y = \frac{U}{W+1}$. Now when $X = 0$ we have $U = 0, W = 0$ when $X \to \infty$ $U \to \infty, V \to \infty$. So we have $U > 0$ and $W > 0$.

Now compute the Jacobian as follows,

$$J = \begin{bmatrix} \frac{w}{1+w} & \frac{-uw}{(1+w)^2} + \frac{u}{1+w} \\ \frac{1}{1+w} & \frac{-u}{(1+w)^2} \end{bmatrix}$$

$$|J| = \frac{u}{(1+w)^2}$$

Since for 2 values of Y we get one value of X we will multiply the jacobian by 2. Now we compute $k(u, w)$ as follows,

$$k(u, w) = f(x, y)|J|$$

$$= \frac{e^{-\frac{uw}{w+1}} \frac{uw}{w+1}^{a-1}}{\Gamma(a)} \times \frac{e^{-\frac{u}{w+1}} \frac{u}{w+1}^{b-1}}{\Gamma(b)} \times \frac{u}{(1+w)^2}$$

$$= \frac{1}{\Gamma(a+b)} e^{-u} u^{a+b-1} \times \frac{1}{\beta(a, b)}$$

Complete this □

Example 1.48. *$X \sim N(\mu, \sigma^2)$. Obtain the distribution of $Y = \frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2$*

Proof. Begin by stating the pdf of r.v. X,

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \qquad -\infty < x < \infty, \sigma > 0$$

$$= 0 \qquad\qquad\qquad \text{otherwise}$$

We now state X in terms of Y as follows, $X = \mu \pm \sqrt{2}\sigma\sqrt{y}$. Range of y is $0 < y < \infty$. And since it is 2-1 correspondence we will multiply the Jacobian by 2.

Compute the value of Jacobian first,

$$|J| = \frac{\sigma}{\sqrt{2}\sqrt{y}}$$

Now compute the new function,

$$h(y) = f(x)|J|2$$

$$= \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \times \frac{\sigma}{\sqrt{2}\sqrt{y}} \times 2$$

$$= \frac{1}{\sigma\sqrt{2\pi}} e^{-y} \frac{2\sigma}{\sqrt{2}\sqrt{y}}$$

$$= \frac{2}{\sqrt{2}\sqrt{y}\sqrt{2\pi}} e^{-y}$$

$$= \frac{e^{-y}}{\sqrt{\pi}\sqrt{y}}$$

$$= \frac{1}{\Gamma(\frac{1}{2})} e^{-y} y^{1-\frac{1}{2}}$$

So we have $Y \sim G\left(\frac{1}{2}\right)$. □

Example 1.49.

$$f(x, y) = \begin{cases} 4xye^{-(x^2+y^2)} & x, y > 0 \\ 0 & otherwise \end{cases}$$

Prove that $h(u) = 2u^3 e^{-u^2}, u > 0$ *where* $u = \sqrt{x^2 + y^2}$ *and* $v = x$.

Proof. The variables we are dealing with are,

$$x = v, y = \sqrt{u^2 - v^2}$$

The range for v, u is $(0, \infty)$ but $0 < v < u < \infty$.
Begin by computing the Jacobian,

$$|J| = \frac{u}{\sqrt{u^2 - v^2}}$$

Consider now the joint distribution with the change of variables,

$$\begin{aligned}
g(u, v) &= f(x, y)|J| \\
&= 4xye^{-(x^2+y^2)}|J| \\
&= 4(v)(\sqrt{u^2 - v^2})e^{-(v^2+u^2-v^2)}\frac{u}{\sqrt{u^2 - v^2}} \\
&= 4v\sqrt{u^2 - v^2}e^{-u^2}\frac{u}{\sqrt{u^2 - v^2}} \\
&= 4vue^{-u^2}
\end{aligned}$$

Integrate out v

$$\begin{aligned}
h(u) &= 4ue^{-u^2}\int_0^u v\,dv \\
&= 4ue^{-u^2}\frac{u^2}{2} \\
&= 2u^3 e^{-u^2}
\end{aligned}$$

$\square$

Example 1.50.

$$f(x, y) = \begin{cases} \frac{e^{-(x+y)}x^3y^4}{\Gamma(4)\Gamma(5)} & x, y > 0 \\ 0 & otherwise \end{cases}$$

Obtain pdf of u where $u = \frac{x}{x+y}$ *take* $v = x + y$ *also obtain* $E[u], V[u]$.

Proof. Consider the new variables, $x = uv, y = v - uv$. The range for v is $(0, \infty)$ and for u is $(0, 1)$

Compute the Jacobian,

$$|J| = v$$

Compute the joint pdf,

$$
\begin{aligned}
g(u, v) &= f(x, y)|J| \\
&= \frac{e^{-(x+y)} x^3 y^4}{\Gamma(4)\Gamma(5)} |J| \\
&= \frac{e^{-(uv+v-uv)} (uv)^3 (v - uv)^4}{\Gamma(4)\Gamma(5)} v \\
&= \frac{e^{-v} u^3 (1 - u)^4 v^8}{\Gamma(4)\Gamma(5)}
\end{aligned}
$$

Integrate out v

$$
\begin{aligned}
h(u) &= \frac{u^3 (1 - u)^4}{\Gamma(4)\Gamma(5)} \int_0^\infty e^{-v} v^8 \, dv \\
&= \frac{\Gamma(9)}{\Gamma(4)\Gamma(5)} u^3 (1 - u)^4
\end{aligned}
$$

So $U \sim \beta_1(m = 4, n = 5)$. Compute the mean and variance as follows,

$$
\begin{aligned}
E[U] &= \frac{m}{m + n} = \frac{4}{9} \\
V[U] &= \frac{mn}{(m + n)^2 (m + n + 1)} = \frac{20}{810} = \frac{2}{81}
\end{aligned}
$$

$\square$

Example 1.51. *X, Y are two independent gamma variates with parameters a, b respectively. Show that $U + X + Y, V = \frac{X-Y}{X+Y}$ are independent.*

Proof. Consider the original pdfs,

$$
\begin{aligned}
f_1(x) &= \frac{e^{-x} x^{a-1}}{\Gamma(a)} \\
f_2(y) &= \frac{e^{-y} x^{b-1}}{\Gamma(b)}
\end{aligned}
$$

Since they are independent

$$f(x,y) = \frac{e^{-(x+y)}x^{a-1}y^{b-1}}{\Gamma(a)\Gamma(b)}$$

Consider the new variables, $x = \frac{1}{2}(uv + u), y = \frac{1}{2}(u - uv)$, the ranges for u is $(0,\infty)$ but for v is $(-1, 1)$
Compute the Jacobian,

$$|J| = \frac{u}{2}$$

Compute the joint pdf,

$$\begin{aligned}
g(u,v) &= f(x,y)|J| \\
&= \frac{e^{-(x+y)}x^{a-1}y^{b-1}}{\Gamma(a)\Gamma(b)}|J| \\
&= -\frac{e^{-u}(v+1)2^{-a-b+2}(u(v+1))^{a-2}(u-uv)^b}{(v-1)\Gamma(a)\Gamma(b)}
\end{aligned}$$

Split this up I'm too lazy to type it.   □

Chapter 2

Chi-square distribution

2.1 PDF

2.2 MGF

2.3 CGF

Theorem 2.1. *The CGF of Chi squared disitributed is given by*

$$K_x(t) = -\frac{n}{2}\log(1 - 2t)$$

Proof. MGF is given by $(1 - 2t)^{-n/2}$ and since CGF is just log MGF. We have the result as required. $\qquad\square$

The values of the cumulants are given as follows,

$$K_1 = n$$
$$K_2 = 2n$$
$$K_3 = 8n$$
$$K_4 = 48n \text{ so } \mu_4 = 48n + 12n^2$$

2.4 Skewness and kurtosis

Theorem 2.2. *Skewness is*

$$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{8}{n}$$

so it is positively skewed.

Theorem 2.3. *Kurtosis is*

$$\beta_2 = \frac{\mu_4}{\mu_2^2} = 3 + \frac{12}{n}$$

so it is leptokurtic and approaches normal as $n \to \infty$.

Plot a graph of chi square with 6 degrees of freedom.

2.5   Additive property of chi squared distribution

Theorem 2.4. *If $X_1, X_2, \ldots, X_k$ are k independent χ^2 variates with degrees of freedom n_i for $i = 1, 2, \ldots, k$ then*

$$\sum_i^n X_i \sim \chi^2_{\sum n_i}$$

Proof. Take the MGF multiply it that's it. □

2.6   Mode

We will obtain the mode as follows,

1. $f'(x) = 0$

2. Check if $f''(x) < 0$

Begin with the pdf,

$$f(x) = \begin{cases} \frac{1}{2^{n/2}\Gamma(n/2)} e^{-x/2} x^{n/2-1} & x > 0, n > 2 \\ 0 & o.w. \end{cases}$$

Now take its logarithm,

$$\log f(x) = \frac{f'(x)}{f(x)}$$

$$\frac{d \log f(x)}{dx} = 0 \implies x = n - 2$$

$$\frac{d^2 \log f(x)}{dx^2} = 0 - \frac{n-2}{2x^2} < 0$$

You can also do it by just taking the derivative and doing it the long way.

You can do it in a third method by getting a recurrence relation after derivative it once.

Theorem 2.5 (Mode of Chi-square distribution)**.** *Mode of Chi-square distribution with n degrees of freedom is $n - 2$*

2.7 Applications of Chi-square distribution

2.7.1 Goodness of fit

This is a very powerful test for testing significance of difference between theoretical and experimental values. It was given by Prof. Karl Pearson in 1900.

If $O_i, i = 1, 2, \ldots, k$ is a set of observed or experimental frequencies and $E_I, i = 1, 2, \ldots, k$ is a set of corresponding expected (hypothetical) frequencies, then Karl Pearson's χ^2 statistic is given by,

H_0 : There is no significant difference between the observed and expected frequencies (the dist. is a good fit).

H_1 : There is significant difference.

The test statistic is given by,

$$\chi^2 = \frac{\sum_{i=1}^{k}(O_i - E_i)^2}{E_i} \sim \chi^2_{k-1, \alpha}$$

The decision criteria is,

Reject H_0 if $\chi^2_{cal} > \chi^2_{tab} = \chi^2_{\alpha, k-1}$

For computation

$$\chi^2 = \frac{\sum O_i^2}{E_i} - N$$

The test is one sided (right sided).

Conditions for validity of the χ^2 test,

1. Sample observations should be independent.

2. Constraints on the cell frequencies should be similar, i.e. $\sum O_i = \sum E_i$.

3. The total frequencies (N) should be reasonably large. Say $N > 50$.

4. No theoretical frequencies should be less than 5. If a frequency is less than 5 then it is pooled with the preceding or succeeding frequency such that the pooled frequency is greater than 5. The degrees of freedom will decrease by the number of observations that are pooled.

Note: Sometimes while fitting a distribution the given data, some parameters have to be estimated. If p parameters are estimated, the degrees of

freedom will be $k-p-1$ where k is the number of classes. and p is the number of parameters estimated. Therefore, $d.f. = (k-1) - (\text{no of pooled values}) - (\text{no. of parameters estimated})$

Example 2.6. *Four identical coins are tossed* 160 *times, and the number of heads is recorded as follows,* X : *No. of heads*

X	0	1	2	3	4
O_i	14	30	70	35	11

Test the hypothesis that the coins are perfect.

Proof. H_0 : Coins are perfect or Binomial is a good fit.
H_1 : Not H_0
   We say X : No. of heads in 4 coin tosses and $X \sim B(n = 4, p = 0.5)$ and $N = 160$.
   We get the following,

X	0	1	2	3	4
O_i	14	30	70	35	11
E_i	10	40	60	40	10

So $\chi^2_{cal} = \frac{\sum O_i^2}{E_i} - N = 6.4917$ and $\chi^2_{tab} = 9.488$. So $\chi^2_{cal} < \chi^2_{tab}$ so it doesn't fall in critical reason and we fail to reject H_0. □

Example 2.7. *A study of printing mistakes in a book of* 550 *pages gives the following data.*
X : *No of printing mistakes per page*

X	0	1	2	3	4
O_i	485	52	8	4	1

Proof. H_0 : Poisson distribution is a good fit
H_1 : Not H_0
$X \sim P(0.15\overline{27} = \frac{42}{275})$
   We get the following,

X	0	1	2	3	4
O_i	485	52	8	4	1
E_i	472	72	6	0	0

Pool the last 3 classes.
   So $\chi^2_{cal} = \frac{\sum O_i^2}{E_i} - N = 14.08027$ and $\chi^2_{tab} = \chi^2_{0.05,1} = 3.841$. So $\chi^2_{cal} > \chi^2_{tab}$ so it falls in critical reason and we accept H_1. □

Chapter 3

F-distribution