VE401 Probabilistic Methods in Eng.
Solution Manual for RC 3

Chen Xiwen

March 21, 2020

University of Michigan–Shanghai Jiao Tong University Joint Institute

Closeness of Binomial and Hypergeometric Distributions

Theorem 1. *Suppose Y has a binomial distribution with parameters $n \in \mathbb{N} \setminus \{0\}$ and $p, 0 < p < 1$. Let $\{X_k\}$ be a sequence of hypergeometric random variables with parameters N_k, n, r_k such that*

$$\lim_{k\to\infty} r_k = \infty, \quad \lim_{k\to\infty} N_k - r_k = \infty, \quad \lim_{k\to\infty} \frac{r_k}{N_k} = p.$$

Then for each fixed n and each $x = 0, \dots, n$,

$$\lim_{k\to\infty} \frac{P[Y = x]}{P[X_k = x]} = 1.$$

Proof. For large enough k, we have

$$\begin{aligned} P[X_k = x] &= \frac{\dfrac{r_k!}{x!(r_k - x)!} \cdot \dfrac{(N_k - r_k)!}{(n-x)!(N_k - r_k - n + x)!}}{\dfrac{N_k!}{n!(N_k - n)!}} \\ &= \binom{n}{x} \frac{r_k!(N_k - r_k)!(N_k - n)!}{(r_k - x)!N_k!(N_k - r_k - n + x)!}. \end{aligned}$$

Recall Stirling's formula from assignment 2,

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n,$$

then we have

$$P[X_k = x] \sim \binom{n}{x} \frac{r_k^{r_k + 1/2}}{(r_k - x)^{r_k - x + 1/2}} \cdot \frac{(N_k - r_k)^{N_k - r_k + 1/2}}{(N_k - r_k - n + x)^{N_k - r_k - n + x + 1/2}} \cdot \frac{(N_k - n)^{N_k - n + 1/2}}{N_k^{N_k + 1/2}}.$$

Then from calculus we have

$$\begin{aligned} \lim_{k\to\infty} \left(\frac{r_k}{r_k - x}\right)^{r_k - x + 1/2} &= e^x, \\ \lim_{k\to\infty} \left(\frac{N_k - r_k}{N_k - r_k - n + x}\right)^{N_k - r_k - n + x + 1/2} &= e^{n-x} \\ \lim_{k\to\infty} \left(\frac{N_k - n}{N_k}\right)^{N_k - n + 1/2} &= e^{-n}. \end{aligned}$$

Therefore,

$$P[X_k = x] \sim \binom{n}{x} \frac{r_k^x (N_k - r_k)^{n-x}}{N_k^n} \sim \binom{n}{x} p^x (1-p)^{n-x} = P[Y = x].$$

$\square$

Exercise 1.

Suppose Y is the rate (calls per hour) at which calls arrive at a switchboard. Let X be the number of calls during a two-hour period. Suppose the joint probability density function is given by

$$f_{XY}(x,y) = \begin{cases} \dfrac{(2y)^x}{x!}e^{-3y} & \text{for } y > 0 \text{ and } x = 0, 1, \dots, \\ 0 & \text{otherwise.} \end{cases}$$

1. Verify that f is a proper joint probability density function.
2. Find $P[X = 0]$.

Solution.

1. To verify that f is a proper joint probability density function, we have

$$\begin{aligned} \int_0^\infty \left(\sum_{x=0}^\infty f_{XY}(x,y) \right) \mathrm{d}y &= \int_0^\infty \left(\sum_{x=0}^\infty \frac{(2y)^x}{x!} \right) e^{-3y} \mathrm{d}y \\ &= \int_0^\infty e^{-y} \mathrm{dy} \\ &= -e^{-y} \Big|_0^\infty = 1. \end{aligned}$$

2. Plugging in $x = 0$ and integrating with respect to y,

$$P[X = 0] = \int_0^\infty f_{XY}(0,y)\mathrm{d}y = \int_0^\infty e^{-3y}\mathrm{d}y = \frac{1}{3}.$$

Exercise 2.

Suppose that X_1 and X_2 are independent random variables, so that

$$X_1 \sim \mathrm{B}(n_1, p), \qquad X_2 \sim \mathrm{B}(n_2, p).$$

For each fixed value of $k (k = 1, 2, \dots, n_1 + n_2)$, prove that the conditional distribution of X_1 given that $X_1 + X_2 = k$ is hyper-geometric with parameters $n_1 + n_2, k, n_1$.

Solution. For $x = 1, \dots, k$,

$$P[X_1 = x | X_1 + X_2 = k] = \frac{P[X_1 = x \text{ and } X_1 + X_2 = k]}{P[X_1 + X_2 = k]} = \frac{P[X_1 = x \text{ and } X_2 = k - x]}{P[X_1 + X_2 = k]}.$$

Since X_1 and X_2 are independent,

$$P[X_1 = x \text{ and } X_2 = k - x] = P[X_1 = x]P[X_2 = k - x].$$

Furthermore, since X_1 and X_2 follow binomial distributions, the sum of them also follows the binomial distribution with parameters $n_1 + n_2$ and p. Therefore,

$$\begin{aligned}
P[X_1 = x] &= \binom{n_1}{x} p^x (1-p)^{n_1 - x}, \\
P[X_2 = k - x] &= \binom{n_2}{k-x} p^{k-x} (1-p)^{n_2 - k + x}, \\
P[X_1 + X_2 = k] &= \binom{n_1 + n_2}{k} p^k (1-p)^{n_1 + n_2 - k}.
\end{aligned}$$

Thus,

$$P[X_1 = x | X_1 + X_2 = k] = \frac{\binom{n_1}{x}\binom{n_2}{k-x}}{\binom{n_1+n_2}{k}},$$

indicating a hypergeometric distribution with parameters $n_1 + n_2, k, n_1$.

A Second Look into Connections of Distributions

The following two theorems focus on discrete random variables, but hold also in continuous case, by switching sums to integrals in the proofs.

Theorem 2. *Suppose $X_1, \ldots, X_n$ are independent discrete random variables, then*

$$\mathrm{E}\left[\prod_{i=1}^n X_i\right] = \prod_{i=1}^n \mathrm{E}\left[X_i\right].$$

Note. *This also holds for* $\mathrm{E}[\varphi \circ X]$.

Proof. By definition, the we calculate the expectation using joint probability density function

$$\begin{aligned}
\mathrm{E}\left[\prod_{i=1}^n X_i\right] &= \sum_{x_1,\ldots,x_n} \left(\prod_{i=1}^n x_i\right) f_{\mathbf{X}}(x_1, \ldots, x_n) \\
&= \sum_{x_1,\ldots,x_n} \left(\prod_{i=1}^n x_i\right) f_{X_1}(x_1) \cdots f_{X_n}(x_n) \\
&= \prod_{i=1}^n \sum_{x_i} x_i f_{X_i}(x_i) = \prod_{i=1}^n \mathrm{E}\left[X_i\right].
\end{aligned}$$

□

Theorem 3. *Suppose $X_1, \ldots, X_n$ are n independent random variables, and for $i = 1, \ldots, n$, let m_{X_i} denote the m.g.f. of X_i. Then the random variable*

$$X = X_1 + \cdots + X_n$$

has the moment generating function

$$m_X(t) = \prod_{i=1}^{n} m_{X_i}(t),$$

for every t such that $m_{X_i}(t)$ is finite for $i = 1, \ldots, n$.

Proof. By definition,

$$m_X(t) = \mathrm{E}\left[e^{tX}\right] = \mathrm{E}\left[e^{t(X_1+\cdots+X_n)}\right] = \mathrm{E}\left[\prod_{i=1}^{n} e^{tX_i}\right].$$

Since $X_1, \ldots, X_n$ are independent, if follows from Theorem 2 that

$$\mathrm{E}\left[\prod_{i=1}^{n} e^{tX_i}\right] = \prod_{i=1}^{n} \mathrm{E}\left[e^{tX_i}\right],$$

and thus

$$m_X(t) = \prod_{i=1}^{n} m_{X_i}(t).$$

□

Sum of Independent Discrete Random Variables

Theorem 4. *Let X and Y be two independent integer-valued random variables, with probability density functions $f_X(x)$ and $f_Y(y)$ respectively. Then the density function of the sum of the random variables $Z = X + Y$ is given by*

$$f_Z(z) = (f_X * f_Y)(z) = \sum_k f_X(k) \cdot f_Y(z-k),$$

where $*$ is the discrete convolution operation. For the sum of independent random variables $S_n = X_1, \ldots, X_n$, we write as

$$S_n = S_{n-1} + X_n,$$

and calculate by induction.

Proof. Using the law of total probability,

$$\begin{aligned}
P[Z=z] &= \sum_x P[Z=z|X=x]P[X=x] \\
&= \sum_x P[Y=z-x|X=x]P[X=x] \\
&= \sum_x P[Y=z-x]P[X=x] \qquad \text{(independence)} \\
&= \sum_k f_X(k)\cdot f_Y(z-k).
\end{aligned}$$

□

Example — Sum of Independent Poisson Distributions

Suppose $X_1,\ldots,X_n$ are independent Poisson distributions with parameters $k_1,\ldots,k_n$. Then the sum of these random variables $X = X_1+\cdots+X_n$ follows the Poisson distribution with parameter $k = k_1+\cdots+k_n$.

Proof. Following Theorem 4 or using m.g.f., we have the following two methods.

- Induction method. Denote

$$S_n = \sum_{i=1}^n X_i,$$

and then

$$S_2 = X_1 + X_2.$$

Knowing the probability density function for X_i, we have

$$f_{X_i}(x) = \frac{k_i^x e^{-k_i}}{x!}, \qquad x\in\mathbb{N}.$$

Therefore,

$$\begin{aligned}
f_{S_2}(s) &= \sum_{x=0}^s \frac{k_1^x e^{-k_1}}{x!}\cdot\frac{k_2^{s-x}e^{-k_2}}{(s-x)!} \\
&= \sum_{x=0}^s \binom{s}{x}\cdot\frac{k_1^x k_2^{s-x} e^{-(k_1+k_2)}}{s!} \\
&= \frac{e^{-(k_1+k_2)}}{s!}\sum_{x=0}^s\binom{s}{x}k_1^x k_2^{s-x} \\
&= \frac{(k_1+k_2)^s e^{-(k_1+k_2)}}{s!},
\end{aligned}$$

indicating a Poisson distribution with parameter k_1, k_2. Using induction with

$$X = S_n = S_{n-1} + X_n,$$

we can conclude that X follows the Poisson distribution with parameter $k = k_1 + \cdots + k_n$.
Note. Induction steps are necessary if you are doing homework or exam, but are omitted here...

- M.G.F. The m.g.f. of each X_i is given by

$$m_{X_i} : \mathbb{R} \to \mathbb{R}, \qquad m_{X_i}(t) = e^{k_i(e^t - 1)}.$$

Therefore, by Theorem 3, we have the m.g.f. for X

$$m_X : \mathbb{R} \to \mathbb{R}, \qquad m_X(t) = \prod_{i=1}^{n} e^{k_i(e^t-1)} = \exp\left((e^t - 1) \sum_{i=1}^{n} k_i \right).$$

Since m.g.f. is unique, we know that X follows the Poisson distribution with parameter $k = k_1 + \cdots + k_n$.

□

Sum of Independent Continuous Random Variables

Theorem 5. *Let X and Y be two continuous random variables with probability density functions $f_X(x)$ and $f_Y(y)$, respectively. Both density function are defined on $\mathbb{R}$. Then the probability density function of the sum $Z = X + Y$ is given by*

$$f_Z(z) = (f_X * f_Y)(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z - x) \mathrm{d}x.$$

where $*$ is the convolution of continuous functions. For the sum of independent random variables $S_n = X_1, \ldots, X_n$, we write as

$$S_n = S_{n-1} + X_n,$$

and calculate by induction.

Proof. Using transformation of random variables, suppose that $U = Z = X + Y, V = X$, then the transformation

$$H : (X, Y) \mapsto (U, V), \qquad H(x, y) = \begin{pmatrix} x + y \\ x \end{pmatrix},$$

and thus

$$H^{-1}(u, v) = \begin{pmatrix} v \\ u - v \end{pmatrix}.$$

The Jacobian is given by

$$DH^{-1}(u,v) = \begin{pmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix} \quad \Rightarrow \quad \det(DH^{-1}) = -1.$$

Therefore,

$$\begin{aligned} f_{UV}(u,v) = f_{XY}(v, u-v)|\det(DH^{-1})| \quad \Rightarrow \quad f_U(u) &= \int_{-\infty}^{\infty} f_{UV}(u,v)\mathrm{d}v \\ &= \int_{-\infty}^{\infty} f_{XY}(v,u-v)\mathrm{d}v \\ &= \int_{-\infty}^{\infty} f_X(v) f_Y(u-v)\mathrm{d}v, \end{aligned}$$

by independence. Replacing U with Z, we have

$$f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)\mathrm{d}x.$$

□

Example — Sum of Independent Gamma Distributions

Suppose random variables $X_1, \ldots, X_n$ are independent, and each X_i follows the gamma distribution with parameters α_i and β. Then the sum $X = X_1 + \cdots + X_n$ follows the gamma distribution with parameters $\alpha_1 + \cdots + \alpha_n$ and β.
Note. This indicates also that the sum of independent exponential distributions is the gamma distribution, since $\mathrm{Exp}(\beta)$ is equivalent to $\mathrm{Gamma}(1, \beta)$.

Proof. Similar as before, we can use either Theorem 5 or m.g.f.

- Induction method. Denote

$$S_n = \sum_{i=1}^{n} X_i$$

and then

$$S_2 = X_1 + X_2.$$

The probability density function for gamma distribution is given by

$$f_{X_i}(x) = \begin{cases} \dfrac{\beta^{\alpha_i}}{\Gamma(\alpha_i)} x^{\alpha_i - 1} e^{-\beta x} & \text{for } x > 0, \\ 0 & \text{for } x \leq 0. \end{cases}$$

Therefore,

$$\begin{aligned} f_{S_2}(s) &= \int_{-\infty}^{\infty} f_{X_1}(x) f_{X_2}(s-x)\mathrm{d}x \\ &= \int_0^s \frac{\beta^{\alpha_1}}{\Gamma(\alpha_1)} x^{\alpha_1-1} e^{-\beta x} \cdot \frac{\beta^{\alpha_2}}{\Gamma(\alpha_2)} (s-x)^{\alpha_2-1} e^{-\beta(s-x)}\mathrm{d}x \\ &= \beta^{\alpha_1+\alpha_2} e^{-\beta s} \int_0^s \frac{x^{\alpha_1-1}(s-x)^{\alpha_2-1}}{\Gamma(\alpha_1)\Gamma(\alpha_2)}\mathrm{d}x. \end{aligned}$$

Here we use a property of Gamma function as follows.

$$\begin{aligned} \Gamma(x)\Gamma(y) &= \int_0^{\infty} u^{x-1}e^{-u}\mathrm{d}u \cdot \int_0^{\infty} v^{y-1}e^{-v}\mathrm{d}v \\ &= \int_0^{\infty}\int_0^{\infty} u^{x-1}v^{y-1}e^{-(u+v)}\mathrm{d}u\mathrm{d}v. \end{aligned}$$

Substituting

$$u = r\cos^2\theta, \qquad v = r\sin^2\theta,$$

we have

$$J = \begin{pmatrix} \cos^2\theta & -2r\sin\theta\cos\theta \\ \sin^2\theta & 2r\sin\theta\cos\theta \end{pmatrix} \quad \Rightarrow \quad \det(J) = 2r\sin\theta\cos\theta.$$

Therefore,

$$\begin{aligned} \Gamma(x)\Gamma(y) &= 2\int_0^{\infty}\int_0^{\pi/2} r^{x+y-1}e^{-r}\cos^{2x-1}(\theta)\sin^{2y-1}(\theta)\mathrm{d}\theta\mathrm{d}r \\ &= 2\int_0^{\infty} r^{x+y-1}e^{-r}\mathrm{d}r \cdot \int_0^{\pi/2}\cos^{2x-1}(\theta)\sin^{2y-1}(\theta)\mathrm{d}\theta \\ &= \Gamma(x+y)\cdot 2\int_0^{\pi/2}\cos^{2x-1}(\theta)\sin^{2y-1}(\theta)\mathrm{d}\theta \qquad (\text{substitute } t=\cos^2(\theta)) \\ &= \Gamma(x+y)\int_0^1 t^{x-1}(1-t)^{y-1}\mathrm{d}t. \end{aligned}$$

Continuing our proof, we have

$$\begin{aligned} f_{S_2}(s) &= \beta^{\alpha_1+\alpha_2}e^{-\beta s}\int_0^s \frac{x^{\alpha_1-1}(s-x)^{\alpha_2-1}}{\Gamma(\alpha_1)\Gamma(\alpha_2)}\mathrm{d}x \\ &= \frac{\beta^{\alpha_1+\alpha_2}e^{-\beta s}}{\Gamma(\alpha_1)\Gamma(\alpha_2)}\int_0^s \left(\frac{x}{s}\right)^{\alpha_1-1}\left(1-\frac{x}{s}\right)^{\alpha_2-1}\mathrm{d}x\cdot s^{\alpha_1+\alpha_2-2} \qquad (\text{substitute } t=\frac{x}{s}) \\ &= \frac{\beta^{\alpha_1+\alpha_2}e^{-\beta s}}{\Gamma(\alpha_1)\Gamma(\alpha_2)}\int_0^1 t^{\alpha_1-1}(1-t)^{\alpha_2-1}\mathrm{d}t\cdot s^{\alpha_1+\alpha_2-1} \\ &= \frac{\beta^{\alpha_1+\alpha_2}}{\Gamma(\alpha_1+\alpha_2)}s^{\alpha_1+\alpha_2-1}e^{-\beta s}, \end{aligned}$$

indicating a gamma distribution with parameters $\alpha_1 + \alpha_2$ and β. Using induction with

$$X = S_n = S_{n-1} + X_n,$$

we can conclude that X follows the gamma distribution with parameters $\alpha = \alpha_1 + \cdots + \alpha_n$ and β.

Note. AGAIN, induction steps are necessary if you are doing homework or exam, but are omitted here...

- M.G.F. The m.g.f. of each X_i is given by

$$m_{X_i} : (-\infty, \beta) \to \mathbb{R}, \qquad m_{X_i}(t) = \frac{1}{(1 - t/\beta)^{\alpha_i}}.$$

Therefore, by Theorem 3, we have the m.g.f. for X

$$m_X : (-\infty, \beta) \to \mathbb{R}, \qquad m_X(t) = \prod_{i=1}^{n} \frac{1}{(1 - t/\beta)^{\alpha_i}} = \frac{1}{(1 - t/\beta)^{\sum_{i=1}^{n} \alpha_i}},$$

indicating a gamma distribution with parameters $\alpha = \alpha_1 + \cdots + \alpha_n$ and β.

$\square$