L E C T U R E 16

Appendix. GENERALIZED INDEPENDENT TRIALS.

In the foregoing we considered independent trials of a random experiment with just two possible outcomes (Binomial distribution). It is natural to consider next the independent trials of an experiment with several possible outcomes, say r possible outcomes, i. e. in each trial we have A_1, A_2, ..., A_r possible events. We assume that we know nonnegative numbers p_1, p_2, ..., p_r, whose sum is 1 such that at each trial p_k represents the probability that A_k will be the outcome of that trial $\big(P(A_k) = p_k,\ k = 1, ..., r$ and $\sum\limits_{k=1}^{r} p_k = 1\big)$. For Bernoulli repeated trials $r = 2$, $A_1 = A$ (success) and $A_2 = \overline{A}$ (failure) and $p_1 + p_2 = p + 1 - p = 1$. Therefore independent repeated Bernoulli trials is a particular case for $r = 2$. Corresponding to the binomial law, we have the multinomial law: the probability that in n trials the event A_1 will occur k_1 times, the event A_2 will occur k_2 times, ..., the event A_r will occur k_r times, for any nonnegative integers k_j satisfying the condition $k_1 + k_2 + ... + k_r = n$, is given by

$$P_n(k_1, k_2, ..., k_r) = \frac{n!}{k_1!\, k_2! \cdot ... \cdot k_r!} \quad p_1^{k_1} \cdot p_2^{k_2} \cdot ... \cdot p_r^{k_r}. \tag{A1}$$

To prove (A1) one must note only that the number of outcomes in Ω, which contain k_1 A_1's, k_2 A_2's, ..., k_r A_r's, is equal to the number of ways a set of size n can be partitioned into r subsets of sizes k_1, k_2, ..., k_r respectively, which is equal to

$$\frac{n!}{k_1!\, k_2! \cdot ... \cdot k_r!}.$$

Each of these outcomes has probability $p_1^{k_1} \cdot p_2^{k_2} \cdot ... \cdot p_r^{k_r}$. Consequently, (A1) is proved. The name, “multinomial law” derives from the role played by the expression given in the multinomial theorem

$$(a_1 + ... + a_r)^n = \sum_{\substack{0 \le k_i \le n \\ i=1,\dots,r \\ k_1+\dots+k_r=n}} \frac{n!}{k_1!\, k_2! \cdot ... \cdot k_r!} \quad a_1^{k_1} \cdot a_2^{k_2} \cdot ... \cdot a_r^{k_r}.$$

Example 18.5. A bowl contains 10 white, 6 black and 4 red balls. 7 balls are drawn with replacement. What is the probability that

i) exactly 4 white balls were selected;

ii) 4 white, 2 black and 1 red balls were selected.

Solution: i) The required probability, based on the binomial law, is

$$P_7(4) = \binom{7}{4}(0.5)^7 = \frac{35}{512} = 0.068359375;$$

ii) by the multinomial law:

$$n = 7,\ r = 3,\ k_1 = 4,\ k_2 = 2,\ k_3 = 1,\ p_1 = .5,\ p_2 = .3,\ p_3 = .2.$$

That is

$$P_7(4,2,1) = \frac{7!}{4!\,2!\,1!} \quad (0.5)^4 \cdot (0.3)^2 \cdot (0.2)^1 = \frac{189}{1600} = 0.118125.$$

Example 18.6. 10 fair dice are thrown on the smooth surface. What is the probability that

i) exactly four "6" were appeared;

ii) four "6", three "5" and three "4" were appeared.

Solution: i) By the binomial law we obtain the required probability

$$P_n(k) = \binom{n}{k} p^k (1-p)^{n-k}, \qquad k = 0, 1, ..., n. \tag{A2}$$

Exercise (The behavior of the binomial probabilities). Show, that as k goes from 0 to n, probabilities $P_n(k)$ increase monotonically, then decrease monotonically, reaching their largest value when k satisfying the inequalities

$$n \cdot p - (1-p) \leq k \leq n \cdot p + p. \tag{A3}$$

$$P_{10}(4) = \binom{10}{4} \ (1/6)^4 \ (5/6)^6 = \frac{5^7 \cdot 7}{6^9} = 0.0542658758509882;$$

ii) by the multinomial law:

$$n = 10,\ r = 4,\ k_1 = 4,\ k_2 = 3,\ k_3 = 3,\ k_4 = 0,\ p_1 = \frac{1}{6},\ p_2 = \frac{1}{6},\ p_3 = \frac{1}{6},\ p_4 = 0.5.$$

That is

$$P_{10}(4,3,3,0) = \frac{10!}{4!\,3!\,3!\,0!} \quad \left(\frac{1}{6}\right)^4 \cdot \left(\frac{1}{6}\right)^3 \cdot \left(\frac{1}{6}\right)^3 = \frac{700}{10077696} \approx 0.0000694.$$

Example 18.7. An urn contains five white and ten black balls. Eight times in succession a ball is drawn out but it is replaced before the next drawing takes place. What is the probability that the balls drawn were white on two occasions and black on six?

Solution: Since the balls are replaced before the next drawing takes place the condition of the urn is always the same just before every trial, and therefore the chance of drawing a white or a black ball is the same for each of the trials. In other words, the trials are independent. The probability of drawing a white ball is $1/3$ and the probability of drawing a black ball $2/3$. Hence the probability of drawing exactly two white and six black in eight trials is

$$P_8(2) = \binom{8}{2} \cdot \left(\frac{1}{3}\right)^2 \cdot \left(\frac{2}{3}\right)^6 = \frac{1,792}{6,561} = 0.273129.$$

Example 18.8. An urn contains five white and ten black balls. Eight of these are drawn out and placed in another urn. What is the probability that the latter then contains two white and six black balls?

Solution: This example resembles the former one that it might be very simply stated as, What is the probability of drawing exactly two white balls in eight trials? It differs in that the trials are not independent; that is, the probability of drawing a white ball on the first attempt is $\frac{5}{15}$, whereas the probability of drawing a white ball on the second attempt is either $\frac{4}{14}$ or $\frac{5}{14}$ according as the first ball was white or black. We can calculate this probability using classical definition:

$$\frac{\binom{5}{2} \cdot \binom{10}{6}}{\binom{15}{8}}$$

which we can represent in the form

$$\binom{8}{2} \cdot \frac{5}{15}\,\frac{4}{14}\,\frac{10}{13}\,\frac{9}{12}\,\frac{8}{11}\,\frac{7}{10}\,\frac{6}{9}\,\frac{5}{8} = \frac{140}{429} = 0.3263403.$$

Example 18.9. Suppose that 20% of all copies of a particular textbook fail a certain binding strength test. Let $\eta(\omega)$ denote the number among 15 randomly selected copies that fail the test. What is the probability that

a) at most eight fail the test;

b) exactly eight fail;

c) at least eight fail;

d) between 4 and 7, inclusive, fail.

Solution:

a) The probability that at most eight fail the test is

$$P(\eta(\omega) \leq 8) = \sum_{k=0}^{8} P_{15}(k) = \sum_{k=0}^{8} \binom{15}{k} 0.2^k \, 0.8^{15-k} = 0.999.$$

b) The probability that exactly eight fail is

$$P(\eta(\omega) = 8) = \binom{15}{8} 0.2^8 \, 0.8^7 = 0.003.$$

c) The probability that at least eight fail is

$$P(\eta(\omega) \geq 8) = 1 - P(\eta(\omega) \leq 7) = 1 - \sum_{k=0}^{7} P_{15}(k) = 1 - \sum_{k=0}^{7} \binom{15}{k} 0.2^k \, 0.8^{15-k} = 1 - 0.996 = 0.004.$$

d) Finally, the probability that between 4 and 7, inclusive, fail is

$$P(4 \leq \eta(\omega) \leq 7) = \sum_{k=4}^{7} P_{15}(k) = \sum_{k=4}^{7} \binom{15}{k} 0.2^k \, 0.8^{15-k} = 0.348.$$

§19. Poisson-Dirichlet Distribution

Recall also that there is a relation between Beta and Gamma functions:

$$B(a,b) = B(b,a) = \int_0^1 x^{a-1} (1-x)^{b-1} \, dx =$$

(if we make the change of variable $x = \dfrac{y}{1+y}$)

$$= \int_0^{+\infty} \frac{y^{a-1}}{(1+y)^{a+b}} \, dy = \frac{\Gamma(a)\,\Gamma(b)}{\Gamma(a+b)} \tag{19.0}$$

It is often necessary to consider random vectors

$$p = (p_1, p_2, ..., p_n) \tag{19.1}$$

that form a discrete probability distribution, i.e. satisfy the following conditions:

$$p_j \geq 0, \quad j = 1, 2, ..., n, \quad \sum_{j=1}^{n} p_j = 1. \tag{19.2}$$

For example, p_j may specify the jth of n possible species in the biological population. Random probabilistic vectors of this type also often arise in the Bayesian approach to statistics.

The simplest non-trivial example of a probability distribution on a simplex Δ_n of vectors satisfying conditions (19.2) is the Dirichlet distribution $D(\alpha_1, \alpha_2, ..., \alpha_n)$, the density of which (with respect to $(n-1)$-dimensional Lebesgue measure on Δ_n) is given by the formula

$$f(p_1, p_2, ..., p_n) = \frac{\Gamma(\alpha_1 + \alpha_2 + ... + \alpha_n)}{\Gamma(\alpha_1)\,\Gamma(\alpha_2)\,...\,\Gamma(\alpha_n)}\; p_1^{\alpha_1 - 1} p_2^{\alpha_2 - 1} ... p_n^{\alpha_n - 1}. \tag{19.3}$$

Parameters α_j can take any strictly positive values, and the nature of the distribution significantly depends on these parameters. If $\alpha_j = 1$ for all j, then we have a uniform distribution on the simplex Δ_n.

Let us verify that density function $f(p_1, p_2, ..., p_n)$ defines by (19.3) satisfies the condition:

$$\int_{\Delta_n} f(p_1, p_2, ..., p_n)\, dp_1\, dp_2 \,...\, dp_n = 1.$$

Proof. Consider particular case $n = 2$:

$$f(p_1, p_2) = \frac{\Gamma(\alpha_1 + \alpha_2)}{\Gamma(\alpha_1)\,\Gamma(\alpha_2)}\; p_1^{\alpha_1 - 1} p_2^{\alpha_2 - 1}.$$

Since Δ_2 is the following domain: $p_1 \geq 0$, $p_2 \geq 0$ and $p_1 + p_2 = 1$ (isosceles right triangle), we get

$$\frac{\Gamma(\alpha_1 + \alpha_2)}{\Gamma(\alpha_1)\,\Gamma(\alpha_2)} \int_0^1 p_1^{\alpha_1 - 1}(1 - p_1)^{\alpha_2 - 1}\, dp_1 = \frac{\Gamma(\alpha_1 + \alpha_2)}{\Gamma(\alpha_1)\,\Gamma(\alpha_2)}\, B(\alpha_1, \alpha_2) = 1,$$

here we use (19.0).

Now consider particular case $n = 3$:

$$f(p_1, p_2, p_3) = \frac{\Gamma(\alpha_1 + \alpha_2 + \alpha_3)}{\Gamma(\alpha_1)\,\Gamma(\alpha_2)\,\Gamma(\alpha_3)}\; p_1^{\alpha_1 - 1} p_2^{\alpha_2 - 1}\, p_3^{\alpha_3 - 1}.$$

Since Δ_3 is the following domain: $p_1 \geq 0$, $p_2 \geq 0$, $p_3 \geq 0$ and $p_1 + p_2 + p_3 = 1$ (tetrahedron), we get

$$\frac{\Gamma(\alpha_1 + \alpha_2 + \alpha_3)}{\Gamma(\alpha_1)\,\Gamma(\alpha_2)\,\Gamma(\alpha_3)} \int_0^1 p_1^{\alpha_1 - 1}\, dp_1 \int_0^{1-p_1} p_2^{\alpha_2 - 1}\,(1 - p_1 - p_2)^{\alpha_3 - 1} dp_2 =$$

If we make the following change of variable $\dfrac{p_2}{1 - p_1} = x$ and $dp_2 = (1 - p_1)\, dx$

$$= \frac{\Gamma(\alpha_1 + \alpha_2 + \alpha_3)}{\Gamma(\alpha_1)\,\Gamma(\alpha_2)\,\Gamma(\alpha_3)} \int_0^1 p_1^{\alpha_1 - 1}\,(1 - p_1)^{\alpha_2 + \alpha_3 - 1}\, dp_1 \int_0^1 x^{\alpha_2 - 1}\,(1 - x)^{\alpha_3 - 1}\, dx =$$

using (19.0) we obtain

$$= \frac{\Gamma(\alpha_1 + \alpha_2 + \alpha_3)}{\Gamma(\alpha_1)\,\Gamma(\alpha_2)\,\Gamma(\alpha_3)}\, B(\alpha_2, \alpha_3) \int_0^1 p_1^{\alpha_1 - 1}\,(1 - p_1)^{\alpha_2 + \alpha_3 - 1}\, dp_1 =$$

$$= \frac{\Gamma(\alpha_1 + \alpha_2 + \alpha_3)}{\Gamma(\alpha_1)\,\Gamma(\alpha_2)\,\Gamma(\alpha_3)}\, B(\alpha_2, \alpha_3)\, B(\alpha_1, \alpha_2 + \alpha_3) =$$

also using (19.0) we finally get

$$= \frac{\Gamma(\alpha_1 + \alpha_2 + \alpha_3)}{\Gamma(\alpha_1)\,\Gamma(\alpha_2)\,\Gamma(\alpha_3)}\, \frac{\Gamma(\alpha_2)\,\Gamma(\alpha_3)}{\Gamma(\alpha_2 + \alpha_3)}\, \frac{\Gamma(\alpha_1)\,\Gamma(\alpha_2 + \alpha_3)}{\Gamma(\alpha_1 + \alpha_2 + \alpha_3)} = 1$$

In general case, we have

$$\int_{\Delta_n} f(p_1, p_2, ..., p_n)\, dp_1\, dp_2 \ldots dp_n =$$

$$= \frac{\Gamma(\alpha_1 + \alpha_2 + ... + \alpha_n)}{\Gamma(\alpha_1)\,\Gamma(\alpha_2) \ldots \Gamma(\alpha_n)} \int_0^1 p_1^{\alpha_1 - 1}\, dp_1 \int_0^{1-p_1} p_2^{\alpha_2 - 1}\, dp_2 \int_0^{1-p_1-p_2} p_3^{\alpha_3 - 1}\, dp_3 \ldots$$

$$\ldots \int_0^{1-p_1-p_2-\ldots-p_{n-2}} p_{n-1}^{\alpha_{n-1} - 1} \left(1 - \sum_{i=1}^{n-1} p_i\right)^{\alpha_n - 1} dp_{n-1} = 1.$$

Lemma 19.1. The Γ-distribution has the additive property that if η_1 and η_2 are independent Γ random variables with $(\alpha_1, 1)$ and $(\alpha_2, 1)$, then $\eta_1 + \eta_2$ is Γ distribution with $(\alpha_1 + \alpha_2, 1)$ parameters, that is, density function of the sum $\eta_1 + \eta_2$ has the following form:

$$f_{\eta_1 + \eta_2}(x) = \frac{1}{\Gamma(\alpha_1 + \alpha_2)}\; x^{\alpha_1 + \alpha_2 - 1}\, e^{-x}, \qquad \text{for} \quad x > 0,$$

and 0 for $x \leq 0$.

Proof. Recall that Γ random variable with parameters (α, β), $(\alpha > 0,\ \beta > 0)$ has the following density function:

$$f(x, \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)}\, x^{\alpha-1}\, e^{-\beta\, x}, \qquad \text{for} \quad x > 0,$$

and 0 for $x \leq 0$.

We know that the density function $f(x)$ of a sum of two independent random variables with density functions $f_1(x)$ and $f_2(x)$ has the form:

$$f_{\eta_1+\eta_2}(x) = \int_{-\infty}^{+\infty} f_1(y)\, f_2(x-y)\, dy. \tag{19.3}$$

Substituting the form of $f_1(y)$ and $f_2(x-y)$ in (19.3), we obtain

$$f_{\eta_1+\eta_2}(x) = \frac{1}{\Gamma(\alpha_1)\,\Gamma(\alpha_2)} \int_0^x y^{\alpha_1-1}\,(x-y)^{\alpha_2-1}\, e^{-y}\, e^{-(x-y)}\, dy =$$

$$= \frac{1}{\Gamma(\alpha_1)\,\Gamma(\alpha_2)}\, e^{-x} \int_0^x y^{\alpha_1-1}\,(x-y)^{\alpha_2-1}\, dy =$$

$$= \frac{x^{\alpha_2-1}}{\Gamma(\alpha_1)\,\Gamma(\alpha_2)}\, e^{-x} \int_0^x y^{\alpha_1-1} \left(1 - \frac{y}{x}\right)^{\alpha_2-1} dy.$$

Make the change of variable $\dfrac{y}{x} = z$, $dy = z\, dz$, we obtain

$$f_{\eta_1+\eta_2}(x) = \frac{x^{\alpha_1+\alpha_2-1}}{\Gamma(\alpha_1)\,\Gamma(\alpha_2)}\, e^{-x} \int_0^1 z^{\alpha_1-1}\,(1-z)^{\alpha_2-1}\, dz = \frac{x^{\alpha_1+\alpha_2-1}}{\Gamma(\alpha_1)\,\Gamma(\alpha_2)}\, e^{-x}\, B(\alpha_1, \alpha_2). \tag{19.4}$$

Therefore substituting (19.0) into (19.4) we obtain

$$f_{\eta_1+\eta_2}(x) = \frac{x^{\alpha_1+\alpha_2-1}}{\Gamma(\alpha_1)\,\Gamma(\alpha_2)}\, e^{-x}\, \frac{\Gamma(\alpha_1)\,\Gamma(\alpha_2)}{\Gamma(\alpha_1+\alpha_2)} = \frac{x^{\alpha_1+\alpha_2-1}}{\Gamma(\alpha_1+\alpha_2)}\, e^{-x}.$$

The proof is complete.