Grabner, Knopfmacher, and Prodinger [1] considered runs in samples of iid Geom(p) random variables $(\Gamma_1, \dots, \Gamma_n)$ (Γ is Geom(p) if $\mathsf{Pr}(\Gamma = j) = q^{j-1}p$, $j = 1, 2, \dots,$ where $q = 1 - p$). A run is a succession of the same symbols: for example $(2, 1, 1, 2, 2, 2, 5, 3, 3, 3)$ has 5 runs (of lengths: 1, 2, 3, 1, and 3). They derived, among other things, the expressions for the expected value, the variance, and the limiting distribution of the number of runs in such samples (see, Proposition 1, Proposition 2, and Theorem 2, respectively in [1]). Their arguments are based on generating functions. Here is a different take (based on CLT for m–dependent random variables [2]):

(i) Show that the number of runs is equal to R_n which is defined by

$$R_n = 1 + \sum_{j=1}^{n-1} X_j, \quad \text{where} \quad X_j := I_{\Gamma_{j+1} \neq \Gamma_j}.$$

(ii) Show that (X_j), $j = 1, \dots, n-1$ are identically distributed 1-dependent random variables (see e. g. [2] for the definition).

(iii) Show that

$$\mathsf{E}R_n = 1 + (n-1)\mathsf{Pr}(\Gamma_2 \neq \Gamma_1) = 1 + (n-1)\frac{2q}{1+q} = \frac{2q}{1+q}n + \frac{1-q}{1+q},$$

and that

$$\operatorname{var}(R_n) = \frac{2q(1-q)^2(2+q^2)}{(1+q)^2(1-q^3)}n - \frac{2q(1-q^2)(3-q+q^2)}{(1+q)^2(1-q^3)}.$$

(iv) Conclude by applying the Hoeffding–Robbins CLT.

REFERENCES

[1] A. Grabner, P. Knopfmacher and H. Prodinger. Combinatorics of geometrically distributed random variables: Run statistics. *LATIN 2000, Lecture Notes in Computer Science 1776 (2000)*, pages 457–462, 2000.

[2] W. Hoeffding and H. Robbins. The central limit theorem for dependent random variables. *Duke Math. J.*, 15:773–780, 1948.